

**UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO DE CIÊNCIAS HUMANAS E NATURAIS
PROGRAMA DE PÓS-GRADUAÇÃO *STRICTO SENSU* EM
LINGUÍSTICA**

CARLOS EDUARDO DEOCLÉCIO

**VARIAÇÃO SINTÁTICA DAS ORAÇÕES ADVERBIAIS
FINAIS: SIMILARIDADES E DIFERENÇAS ENTRE FALA E
ESCRITA**

VITÓRIA

2011

CARLOS EDUARDO DEOCLÉCIO

**VARIAÇÃO SINTÁTICA DAS ORAÇÕES ADVERBIAIS FINAIS:
SIMILARIDADES E DIFERENÇAS ENTRE FALA E ESCRITA**

Dissertação apresentada ao Programa de Pós-graduação *stricto sensu* em Linguística do Centro de Ciências Humanas e Naturais da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do título de Mestre em Linguística, na área de concentração Estudos analítico-descritivos da linguagem.

Orientadora: Prof.^a Dr.^a Maria Marta Pereira Scherre

VITÓRIA

2011

Dados Internacionais de Catalogação-na-publicação (CIP)
(Biblioteca Central da Universidade Federal do Espírito Santo, ES, Brasil)

D295v Deoclécio, Carlos Eduardo, 1981-
Variação sintática das orações adverbiais finais :
similaridades e diferenças entre fala e escrita / Carlos Eduardo
Deoclécio. – 2011.
124 f.

Orientador: Maria Marta Pereira Scherre.
Dissertação (Mestrado em Estudos Linguísticos) –
Universidade Federal do Espírito Santo, Centro de Ciências
Humanas e Naturais.

1. Fala. 2. Escrita. 3. Sociolinguística. 4. Variação sintática. I.
Scherre, Maria Marta Pereira, 1950-. II. Universidade Federal do
Espírito Santo. Centro de Ciências Humanas e Naturais. III.
Título.

CDU: 80

CARLOS EDUARDO DEOCLÉCIO

**VARIAÇÃO SINTÁTICA DAS ORAÇÕES ADVERBIAIS FINAIS: SIMILARIDADES
E DIFERENÇAS ENTRE FALA E ESCRITA**

Dissertação apresentada ao Programa de Pós-graduação *stricto sensu* em Linguística do Centro de Ciências Humanas e Naturais da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do Grau de Mestre em Linguística.

COMISSÃO EXAMINADORA

Prof^a. Dr^a. Maria Marta Pereira Scherre - Ufes
Orientadora, Presidente da Sessão e da Comissão Examinadora

Prof^a. Dr^a. Edair Maria Görski - UFSC
Membro Titular Externo da Comissão Examinadora

Prof^a. Dr^a. Lilian Coutinho Yacovenco - Ufes
Membro Titular Interno da Comissão Examinadora

Prof^a. Dr^a. Lúcia Helena Peyroton da Rocha - Ufes
Membro Titular Interno da Comissão Examinadora (Suplente)

À minha mãe, Ana, pelo amor, cuidado e incentivo de sempre; e à minha outra mãe, Angela, que participou do início dessa jornada, mas que está vendo o término dela de algum lugar especial.

AGRADECIMENTOS

À minha família, pelo apoio que sempre deu aos meus estudos, e por compreender as ausências mais do que constantes, particularmente nos últimos meses.

À Prof.^a Maria Marta Pereira Scherre, por ter me recebido como orientando após um ano de início do curso, pela dedicação esmerada, tanto nas aulas, quanto na orientação, pela paixão como conduz seu trabalho, levando-nos a querer fazer sempre melhor, e, finalmente, pela paciência, regada por seu bom humor.

À Prof.^a Lúcia Helena Peyroton da Rocha, que me apoiou na fase inicial do curso, como professora e primeira orientadora, e compreendeu as mudanças de rumo durante o mestrado.

À Prof.^a Lilian Coutinho Yacovenço, pela atenção como professora nas disciplinas cursadas, no mestrado e na graduação, e por aceitar compor a banca de defesa.

À Prof.^a Edair Maria Görski, pelo aceite imediato na composição da banca, bem como pelo fornecimento de material de pesquisa.

À Prof. Hilda de Oliveira Olímpio, que participou da banca de qualificação, trazendo ótimas contribuições ao texto e às questões da pesquisa, e também por ter participado de boa parte do meu processo de formação no curso de Letras na Ufes.

Aos amigos Hosanna e Fausto, que me apoiaram imensamente durante todo o curso e souberam compreender os meus “nãos”.

Aos colegas do curso, em especial à Kelly e à Gabriela, pelas palavras de incentivo.

Aos colegas do Centro de Línguas para a Comunidade – CLC/Ufes, pelo apoio dado durante a primeira fase do curso.

Aos colegas do Instituto Federal do Espírito Santo – Ifes, pelo apoio recebido na fase de conclusão deste trabalho.

“Paciência e perseverança têm o efeito mágico de fazer as dificuldades desaparecerem e os obstáculos sumirem.”

John Quincy Adams

RESUMO

Esta dissertação tem por finalidade analisar a variação sintática das orações subordinadas adverbiais finais, ora desenvolvidas, ora reduzidas, na fala e na escrita, cujas ocorrências se ilustram nestes exemplos: “[...] a mãe corta essa ligação *para que o filhote aprenda a cuidar de sua vida*” (dado de final desenvolvida, do *corpus* da escrita) e “[...] cheguei em casa chorando pra não... *pra ela não me bater*” (dado de final reduzida, do *corpus* da fala). Para essa análise, parte-se da orientação teórico-metodológica da Sociolinguística Variacionista, sobretudo em Labov (2008), complementada pelo uso de princípios do Funcionalismo norte-americano, como o da marcação (GIVÓN, 1995), o da economia e o da iconicidade (HAIMAN, 1983). O *corpus* da língua falada é formado por 19 células da amostra PORTVIX (conjunto de entrevistas nos moldes labovianos, realizadas com falantes da cidade de Vitória/ES), e o da língua escrita é constituído de 35 reportagens da revista Superinteressante. Para a quantificação dos dados extraídos desses *corpora*, foi utilizado o programa computacional Varbrul, em sua versão GoldVarb X (SANKOFF; SMITH; TAGLIAMONTE, 2005), que propicia uma análise em termos da relevância estatística das variáveis que exercem influência sobre a variação do fenômeno investigado. Em termos globais, o percentual de finais desenvolvidas foi de 3% na fala e 16% na escrita, logo 97% e 84% de reduzidas, respectivamente. Foram controladas cinco variáveis independentes de natureza linguística e três de caráter social, entre as quais foram selecionadas pelo programa, por sua significância estatística, as seguintes: para os dados da fala, *características sintático-semânticas do sujeito da oração principal* (com o fator *sujeito não controlador* favorecendo as finais desenvolvidas, com 23.1% das ocorrências, e peso relativo de 0.94); *explicitação do sujeito da adverbial final* (com o fator *sujeito explícito* aumentando a média das desenvolvidas para 10.8%, com peso relativo de 0.91); e *nível de escolarização* do informante (favorecendo as finais desenvolvidas o fator *ensino superior*, com 6.7% de frequência e peso relativo de 0.84); para os dados da escrita, a *correferencialidade do sujeito da adverbial final*, de maneira que as finais desenvolvidas são favorecidas pelos fatores *sujeito parcialmente correferente* (29.6% de frequência e 0.81 de peso relativo) e *sujeito não correferente* (51.6% de uso e 0.76 de peso relativo); e a *explicitação do sujeito da adverbial final*, em que o fator *sujeito explícito* é favorecedor dessa variante (85.7% de frequência e peso relativo de 0.97). Entende-se, neste trabalho, que a correlação entre a (não) correferencialidade e a (não) explicitação do sujeito da adverbial final ocorre por motivação econômica, em que a primeira motiva a segunda, e que a maior ou menor codificação na posição de sujeito motiva, por iconicidade, a ocorrência da final desenvolvida, no primeiro caso, e a da reduzida no segundo. Depois de aferidos e analisados os resultados obtidos, conclui-se que a variação sintática das adverbiais finais não é um fenômeno de caráter regional, por estar alinhado aos resultados encontrados em trabalhos desenvolvidos previamente, como os de Finck (2000), para a fala, e os de Azevedo (2000), para a escrita. Trata-se, portanto, de um fenômeno substancialmente interno do sistema linguístico do português.

Palavras-chave: variação sintática, orações finais, sociolinguística variacionista, fala, escrita.

ABSTRACT

The aim of this study is to analyze the syntactic variations of final adverbial subordinate clauses, one moment developed, the next reduced, in speech or writing, illustrated in the following examples: “[...] *a mãe corta essa ligação para que o filhote aprenda a cuidar de sua vida*” (example of final developed clause, from the writing corpus) and “[...] *cheguei em casa chorando pra não... pra ela não me bater*” (example of final reduced clause, from the speech corpus). This analysis is based on the theoretical-methodological orientation from Variationist Sociolinguistics, mainly in Labov (2008), complemented by the use of the principles from North American Functionalism, such as marking (GIVÓN, 1995), economy and iconicity (HAIMAN, 1983). The speech language corpus was composed by 19 interviews from PORTVIX sample (group of interviews held with speakers in the city of Vitória/ES, using the Labovian model), whereas the written language corpus was composed by 35 articles from *Superinteressante* magazine. The computer program Varbul, version GOLDBARB X (SANKOFF; SMITH; TAGLIAMONTE, 2005), was used to quantify the data obtained in this corpora. This allowed analyzing the data in terms of statistic relevance of the variables which influence the variation of the investigated phenomenon. Overall, the percentage of final developed clauses was 3% in speech and 16% in writing, thus 97% and 84%, respectively. Five independent variables of linguistic nature and three of social character were identified in the study. By meanings of statistics, the program selected the following: for speech data, *syntactic-semantic characteristics of the subject in the main clause* (with the element of *non controlling subject* benefitting the final developed clause, occurring 23.1%, relatively 0.94); for writing data, explicitness of the subject in the final adverbial clause (with the element of the explicit subject increasing the final developed ratio to 10.8%, relative weight 0.91); and the informer’s education level (with the element of higher education benefitting the final developed clause, with a frequency of 6.7%, relative weight 0.84); for writing data, the co-referentiability of the subject in the final adverbial, in a sense that final developed are benefitted from the elements of partially co-referential subject (with frequency of 29.6%, relatively 0.81) and non co-referential subject (51.6% of use, relative weight 0.76); and the subject explicitness in the final adverbial clause, in which the element of explicit subject is benefitted from this variation (85.7% of use, relative weight 0.97). In this paper it is understood that the correlation between the (non) co-referentiality and the (non) explicitness of the subject in the final adverbial clauses occurs by economic motivation, in which the former motivates the latter. In addition, the more or less coding of the subject motivates the occurrence of the final developed clause in the first case and the final reduced clause in the second by iconicity. After the results were assessed and analyzed, it was concluded that the syntactic variation of the final adverbial clauses is not a regional phenomenon, for it is aligned to the results found in previously developed studies, such as Finck’s (2000) on speech and Azevedo’s (2000) on writing. Hence, it may be considered a phenomenon which is substantially domestic to the linguistic system of the Portuguese language.

Keywords: syntactic variation, final clauses, Variationist Sociolinguistics, speech, writing.

LISTA DE QUADROS

Quadro 1 – Distribuição social dos informantes do <i>corpus</i> da fala	66
Quadro 2 – Distribuição dos jornalistas que distribuída nos grupos de fatores sociais.....	67
Quadro 3 – Aplicação do princípio da marcação às orações finais, com base na associação entre correferencialidade e explicitação do sujeito da adverbial final, na fala e na escrita	111

LISTA DE TABELAS

Tabela 1 – Quantificação das finais desenvolvidas e reduzidas nos dados da fala e da escrita.....	68
Tabela 2 – Quantificação das finais desenvolvidas e reduzidas nos dados da fala: Vitória (PORTVIX) <i>versus</i> Florianópolis (VARISUL)	69
Tabela 3 – Quantificação das finais desenvolvidas e reduzidas nos dados da escrita: Superinteressante <i>versus</i> Azevedo (2000)	70
Tabela 4 – Frequências absoluta e relativa dos dados da fala e da escrita, com base no grupo de fatores correferencialidade do sujeito da adverbial final	71
Tabela 5 – Frequências absoluta e relativa dos dados da fala e da escrita, com base no grupo de fatores características sintático-semânticas do sujeito da oração principal	78
Tabela 6 – Frequências absoluta e relativa dos dados da fala e escrita, com base no grupo de fatores voz/aspecto semântico do verbo da adverbial final	80
Tabela 7 – Frequências absoluta e relativa dos dados da fala e da escrita, com base no grupo de fatores explicitação do sujeito da adverbial final.....	83
Tabela 8 – Frequências absoluta e relativa dos dados da fala e da escrita, com base no grupo de fatores posição da adverbial final	85
Tabela 9 – Frequências absoluta e relativa dos dados da fala e da escrita, com base no grupo de fatores sexo/gênero	90
Tabela 10 – Frequências absoluta e relativa dos dados da fala e da escrita, com base no grupo de fatores faixa etária.....	92
Tabela 11 – Frequências absoluta e relativa dos dados da fala, com base no grupo de fatores nível de escolarização	93
Tabela 12 – Frequências absoluta e relativa de todos os dados, com base no grupo de fatores modalidade de expressão da língua	95

Tabela 13 – Atuação das variáveis selecionadas por relevância estatística no uso das finais desenvolvidas nos dados da fala	98
Tabela 14 – Atuação das variáveis selecionadas por relevância estatística no uso das finais desenvolvidas nos dados da escrita.....	100
Tabela 15 – Cruzamento entre (1) correferencialidade e (4) explicitação do sujeito da oração final.....	102
Tabela 16 – Cruzamento entre (1) correferencialidade e (4) explicitação do sujeito nas finais reduzidas nos dados da fala: Vitória (ES) <i>versus</i> Florianópolis (SC)	107
Tabela 17 – Cruzamento entre (1) correferencialidade e (4) explicitação do sujeito da oração final.....	110
Tabela 18 – Atuação das variáveis selecionadas por relevância estatística no uso de sujeito explícito em orações finais na fala.....	115
Tabela 19 – Síntese dos resultados aferidos para as variáveis independentes e sua influência no uso das finais desenvolvidas nos dados da fala e da escrita	117

SUMÁRIO

INTRODUÇÃO	14
CAPÍTULO 1: AS ORAÇÕES FINAIS E AS PERSPECTIVAS DA PESQUISA	18
1.1 A DESCRIÇÃO DO FENÔMENO LINGÜÍSTICO	18
1.2 POR QUÊ? PARA QUÊ? O QUE SE PROJETA?	25
CAPÍTULO 2: AS ORAÇÕES FINAIS NAS GRAMÁTICAS NORMATIVAS E DESCRITIVAS.....	33
CAPÍTULO 3: FUNDAMENTAÇÃO TEÓRICA E REVISÃO BIBLIOGRÁFICA	41
3.1 PRESSUPOSTOS TEÓRICOS DA SOCIOLINGÜÍSTICA VARIACIONISTA E DO FUNCIONALISMO NORTE-AMERICANO	41
3.2. UM VIÉS FUNCIONALISTA: OS PRINCÍPIOS DA MARCAÇÃO, ECONOMIA E ICONICIDADE.....	45
3.2.1 Marcação	45
3.2.2 Economia e iconicidade.....	50
3.3 REVISÃO BIBLIOGRÁFICA	54
CAPÍTULO 4: METODOLOGIA	57
4.1 OS <i>CORPORA</i> EM ANÁLISE E A SELEÇÃO DOS DADOS.....	57
4.2 O VARBRUL E AS VARIÁVEIS INDEPENDENTES.....	60
CAPÍTULO 5: APRESENTAÇÃO E ANÁLISE DOS DADOS	68
5.1 RESULTADOS DOS GRUPOS DE FATORES DE NATUREZA LINGÜÍSTICA	71
5.1.1 Correferencialidade do sujeito da adverbial final.....	71
5.1.2 Características sintático-semânticas do sujeito da oração principal	77
5.1.3 Voz/aspecto semântico do verbo da adverbial final	80
5.1.4 Explicitação do sujeito da adverbial final	82
5.1.5 Posição da adverbial final.....	85
5.2 RESULTADOS DOS GRUPOS DE FATORES DE NATUREZA SOCIAL	89
5.2.1 Sexo/gênero	90

5.2.2 Faixa etária	92
5.2.3 Nível de escolarização	93
5.3 GRUPOS DE FATORES RELEVANTES ESTATISTICAMENTE	96
5.4 CONSIDERAÇÕES SOBRE A CORREFERENCIALIDADE E A EXPLICITAÇÃO DO SUJEITO DA ADVERBIAL FINAL	101
5.5 UMA ESCALA PARA AS ORAÇÕES ADVERBIAIS FINAIS	109
5.6 A EXPLICITAÇÃO DO SUJEITO COMO VARIÁVEL DEPENDENTE: BREVES CONSIDERAÇÕES	114
 CONSIDERAÇÕES FINAIS	 116
 REFERÊNCIAS	 120
 ANEXOS	 124

INTRODUÇÃO

As línguas, em geral, têm dispostas em seu construto lexical e gramatical possibilidades várias de expressar um mesmo valor de verdade – não para todos os eventos da comunicação linguística humana, mas para muitos deles. Guardadas as proporções das nuances contextuais, discursivas e situacionais de produção de sentido, o usuário de uma língua, desde que tenha tido acesso a essas maneiras diversas de expressão, pode lançar mão de algumas delas, com ou sem consciência e intenção, nas diferentes cenas em que precisa se comunicar, seja pela oralidade, seja pela modalidade escrita da língua – ambas em seus vários gêneros discursivos.

Pela orientação teórica da Sociolinguística Variacionista (LABOV, 2008), entende-se que, em muitos casos, esse “lançar mão” não é uma escolha aleatória, mas que há fatores, de ordem linguística ou social, que condicionam e governam a realização de uma variante linguística em detrimento de outra ou outras.

Nesse sentido, compreender e aceitar que as línguas apresentam uma gama múltipla de formas e funções linguístico-comunicativas torna-se aspecto fundamental para que uma investigação de cunho científico analise fenômenos linguísticos nos distintos segmentos em que a gramática de uma língua se organiza: a fonética, a fonologia, a morfologia, a sintaxe e a semântica, cada qual em suas muitas e amplas ramificações e interseções.

Particularmente nesta linha de pensamento, tratar de fenômenos variáveis de natureza sintática – objeto de estudo deste trabalho – requer um nível de observação e de percepção de suas ocorrências de forma contextual e direcionada. Em outras palavras, quando se decide estudar um fenômeno sintático sob a perspectiva de que ele apresenta uma configuração variável no sistema linguístico, é necessário que se busquem os contextos da língua em uso em que possa emergir sua ocorrência, para assim ser passível de observação, análise e descrição minimamente plausíveis, com assertivas menos intuitivas e categóricas a seu respeito.

Para Labov (2008), a variação é um movimento próprio dos sistemas linguísticos. Cabe ao linguista competente determinar como dado fenômeno linguístico se configura como variável e perceber as regularidades de seu funcionamento. Assim, para o autor,

A capacidade dos seres humanos de aceitar, preservar e interpretar regras com condicionamentos variáveis é sem dúvida um aspecto importante de sua competência linguística ou *langue*. Mas ninguém tem consciência dessa competência, e não existem julgamentos intuitivos acessíveis para revelá-la a nós. Ao contrário, a percepção ingênua do nosso próprio comportamento e do dos outros é normalmente categórica, e somente o estudo cuidadoso da língua em uso demonstrará a existência dessa capacidade de operar com regras variáveis. (LABOV, 2008, p. 263)¹

Sendo assim, para que o linguista/pesquisador possa dar conta das explicações mais coerentes para as tendências de uso de uma ou de outra variante de uma variável linguística, bem como dos fatores linguísticos e/ou sociais que interferem nessas ocorrências, é preciso que ele tenha consciência das pressões que as condições diretamente ligadas ao uso corrente da língua podem exercer sobre a compreensão de um fenômeno linguístico em estudo. Incluem-se nessas condições o fato de os dados em análise pertencerem à modalidade oral ou escrita da língua em seus diversos gêneros; a frequência com que essas variantes são veiculadas nas diferentes situações comunicativas; as possíveis e prováveis restrições internas do sistema linguístico; e também os aspectos sociais referentes ao usuário, tais como o nível de escolarização, a idade, o sexo e a classe social a que pertence, entre outros fatores que se julgarem necessários para uma descrição mais acurada e precisa.

Nesse sentido, o presente trabalho, vinculado à linha de pesquisa dos estudos analítico-descritivos da linguagem, tem como proposta investigar o comportamento variável de um fenômeno linguístico de natureza sintática: o desenvolvimento e a redução de orações subordinadas adverbiais finais, em dados da língua falada e escrita, em dados como estes:

- (1) [...] porque sábado, domingo, os pessoal ligam mais cedo PRA VOCÊ ATENDER TRÊS HORAS DA MANHÃ, DUAS HORAS, aí você atende.
- (2) Todo ano, um substituto de Jesus precisava ser morto PARA QUE O SANGUE PUDESSE SER UTILIZADO NO PÃO DA PÁSCOA JUDAICA e trazer a cura.

O exemplo (1) é uma ocorrência da língua falada, cuja estrutura revela em destaque uma oração adverbial final reduzida, ao passo que o exemplo (2) é um dado da escrita que traz em destaque uma final desenvolvida. Na crença de que (1) possa ocorrer na língua como final

¹ A tradução da edição brasileira da obra *Sociolinguistic Patterns* (1972), de William Labov, foi feita por Marcos Bagno, Maria Marta Pereira Scherre e Carolina Rodrigues Cardoso, e publicada pela Parábola Editorial em 2008.

desenvolvida e (2) como final reduzida, este trabalho busca as razões, linguísticas ou não, pelas quais essa alternância possa vir a ocorrer.

Para tanto, as amostras das quais foram coletados os dados que compõem esta pesquisa se dividem em dois grupos: os da língua falada foram extraídos de parte das entrevistas que compõem o *corpus* do projeto PORTVIX – *O português falado na cidade de Vitória* –, cujo acesso foi facilitado pelo próprio Programa de Pós-graduação em Estudos Linguísticos – PPGEL, da Universidade Federal do Espírito Santo (UFES); os da língua escrita foram coletados de reportagens da Revista Superinteressante, publicação da Editora Abril. Mais detalhes sobre os *corpora* selecionados serão apresentados no capítulo 4, dedicado à metodologia da pesquisa.

Como referencial teórico que sustenta a análise e o entendimento do fenômeno de variação sintática das adverbiais finais, esta dissertação traz como centrais as orientações da Sociolinguística Variacionista, cujas ideias principais serão desenvolvidas mais adiante, no capítulo 3, assentadas principalmente em Labov (2008). Como contribuição teórica complementar, são utilizados alguns princípios do Funcionalismo norte-americano, tais como o da marcação, o da economia e o da iconicidade, a partir de Haiman (1983) e Givón (1995, 2001), aspectos que também serão mais bem discutidos no capítulo 3 deste trabalho.

A arquitetura deste texto está disposta da seguinte maneira: no capítulo 1, é apresentado o objeto de estudo desta pesquisa, em sua estrutura, sua variabilidade de ocorrências, de acordo com a noção de valor de verdade das variantes da variável linguística em análise. Também são apresentadas as razões pelas quais a pesquisa foi desenvolvida, os objetivos com ela pretendidos, bem como as principais hipóteses levantadas preliminarmente à análise subsequente.

O capítulo 2 apresenta as considerações de algumas gramáticas normativas e descritivas do português em relação às orações adverbiais finais, muito mais em termos estruturais que propriamente discursivos. Será visto que as descritivas, principalmente a de Neves (2000), abordam o assunto com mais profundidade que as normativas.

Por sua vez, o capítulo 3 apresenta o alinhamento teórico que fundamenta as discussões estabelecidas neste estudo, que são a Sociolinguística Variacionista e alguns princípios do Funcionalismo norte-americano, conforme se antecipou acima. Este mesmo capítulo traz, ainda, breves comentários a respeito de algumas pesquisas desenvolvidas previamente a esta,

em nível de mestrado e doutorado, por apresentarem alguns pontos de interseção com este trabalho, seja por seu objeto de estudo, seja por sua orientação teórico-metodológica.

Na sequência, o capítulo 4 expõe a metodologia do trabalho desenvolvido, enfatizando os *corpora* selecionados para representar as ocorrências da modalidade oral e escrita da língua portuguesa. Além disso, são apresentadas algumas considerações sobre o programa computacional Varbrul, na versão GoldVarb X (SANKOFF; SMITH; TAGLIAMONTE, 2005), utilizado como ferramenta importante para a quantificação dos dados e para o tratamento da relevância estatística dos fatores envolvidos na variação sintática das adverbiais finais. Ainda neste mesmo capítulo, são expostas as variáveis independentes, ou grupos de fatores, de natureza linguística e social, escolhidas para um melhor entendimento do comportamento variável do fenômeno linguístico aqui investigado.

O capítulo 5, por sua vez, apresenta os resultados obtidos por meio das rodadas feitas com os dados linguísticos inseridos no programa Varbrul, bem como a análise desencadeada a partir deles. A tessitura do capítulo assim se organiza: em primeiro lugar, são apresentados e analisados os resultados para todos os grupos de fatores do estudo, independentemente de terem ou não relevância estatística identificada após as rodadas; em segundo lugar, são focalizadas apenas as variáveis com significância estatística. Pela importância que apresentaram ao longo de toda a análise, resolveu-se dedicar neste capítulo uma seção com considerações acerca da correferencialidade e da explicitação do sujeito da adverbial final e suas implicações para a compreensão da variabilidade entre as finais desenvolvidas e reduzidas. Na sequência do capítulo, essas duas variáveis são retomadas para a apresentação de uma escala decrescente das orações finais, com base no princípio funcionalista da marcação (GIVÓN, 1995) – da mais não marcada para a mais marcada, em termos da frequência de uso.

Finalmente, são apresentadas as principais conclusões obtidas ao longo da pesquisa desenvolvida, bem como o apontamento de perspectivas e possibilidades de novos estudos ainda com as orações finais e com outros tipos de orações em termos da variação sintática.

CAPÍTULO 1

AS ORAÇÕES FINAIS E AS PERSPECTIVAS DA PESQUISA

1.1 A descrição do fenômeno linguístico

Na língua portuguesa, para a expressão da finalidade de natureza oracional circunstancial, seu usuário faz uso das tradicionalmente chamadas orações subordinadas adverbiais finais. Para Said Ali (1964, p.143), esse tipo de oração representa “[...] o intento ou propósito a que se dirige ato expresso na oração subordinante”. De acordo com Neves (2000, p. 888), essas orações “[...] se caracterizam semanticamente como expressão da **finalidade**, ou do **propósito** que motiva o evento expresso na **oração principal**”. A autora acrescenta, ainda, que o “contexto mais característico de uma **oração final** é, pois, em ligação com uma **oração principal** que tenha sujeito capaz de exercer controle sobre o evento expresso na **final**”². A partir dessas definições, serão descritas na sequência as características internas desse tipo de estrutura.

Para um maior dinamismo das explanações, as *orações subordinadas adverbiais finais* serão ora chamadas *orações finais*, ora *adverbiais finais*, ou somente *finais*. Assim, as orações finais são introduzidas, em geral, por conectivos específicos e constituídas por tempos e modos verbais que lhes são peculiares. Dessa maneira, as locuções conjuntivas *para que* e *a fim de que*, quando introduzem essas orações, fazem com que suas formas verbais estejam em *modo finito*, mais precisamente no *modo subjuntivo* – o que configura as chamadas *finais desenvolvidas*. De outro modo, quando a introdução da oração final é feita pela preposição *para* ou pela locução prepositiva *a fim de*, a estrutura verbal se apresenta sob a *forma infinitiva*, o que caracteriza a estrutura típica das *finais reduzidas*, em geral classificadas pela tradição gramatical de *orações subordinadas adverbiais finais reduzidas de infinitivo*. Azeredo (2008, p.329) aponta outros conectivos que podem desempenhar o mesmo papel, que são *de maneira que* e *de molde a que*. Entretanto, na prática, como se verá mais adiante nos dados do uso linguístico, os conectivos recorrentes são *para que* e *para*, com uma frequência bem mais ampla do último. Ainda, deve-se destacar que na modalidade oral da língua essas

² Negritos de acordo com o texto original.

formas passam por uma redução de caráter fonético-fonológico e são realizadas predominantemente *pra que e pra*.

A partir dos dados coletados nos *corpora* da língua falada e escrita em uso, será feita uma exposição de exemplos do fenômeno linguístico que é objeto deste estudo. Ademais, para tornar essa exposição mais clara e coerente sob a ótica variacionista, será feita uma proposição de pares mínimos que conservem o máximo de propriedades sintático-semânticas em comum. A justificativa para fazê-lo vem de uma importante e pertinente observação feita por Oliveira (1987), ao referir-se às noções de *contexto* e *valor de verdade*, com base na seguinte passagem de Tarallo (1985, p. 8): “Variantes linguísticas são, portanto, diversas maneiras de se dizer a mesma coisa em um mesmo contexto, e com o mesmo valor de verdade. A um conjunto de variantes dá-se o nome de ‘variável linguística’.”

Labov (2008, p. 221), ao referir-se às formas alternantes de se dizer o mesmo, assim pondera:

Algumas palavras como *carro* e *automóvel* parecem ter os mesmos referentes; outras têm duas pronúncias, como *cantando* e *cantano*. Existem opções sintáticas como *Uma pessoa que eu confio muito* vs. *Uma pessoa em quem eu confio muito* ou *É fácil para ele falar* vs. *Para ele falar é fácil*. Em cada um destes casos, temos o problema de decidir o lugar desta variação na estrutura linguística.³

Com base nessas considerações, deve-se ratificar que, neste trabalho, o que se está chamando de *variável* é o fenômeno de variação sintática entre as orações finais, que podem apresentar-se sob a forma das *variantes*⁴ desenvolvida e reduzida. Voltando ao que Oliveira (1987) preconiza, é fundamental que o que esteja sendo considerado como variável possa ter as suas variantes ocorrendo num mesmo contexto discursivo – entretanto, não se pode perder de vista que o autor também aborda a questão do contexto social. Em relação à noção de valor de verdade, em princípio se estaria trabalhando com a ideia de mesmo significado referencial para as variantes em questão, uma espécie de paráfrase. Essa atitude tem por objetivo a adoção de um critério que considere, numa análise variacionista, casos que possam ser de variabilidade e invariabilidade linguística, já que esta também pode ser importante para uma compreensão mais global de um dado fenômeno linguístico em análise.

³ Observe-se que os exemplos são uma adaptação para o português feita pelos tradutores.

⁴ Mais adiante essa nomenclatura tomará uma conformação complementar: o fenômeno linguístico em estudo será chamado de *variável dependente*, que será analisada mediante a atribuição de *variáveis independentes*.

A partir deste ponto, expõem-se alguns exemplos de orações finais extraídos dos *corpora* que dão suporte empírico a este trabalho. A exposição será feita da seguinte maneira: primeiro, são colocados dois exemplos das variantes em análise que compartilhem de propriedades sintático-semânticas similares; a seguir, é feita uma breve descrição dessas propriedades para que se ateste a sua possível ocorrência variável.

(1) [...] ele criou condições PARA QUE POVOS NÃO-JUDAICOS, ao receberem a mensagem de Deus, FOSSEM INSERIDOS DE FORMA IGUALITÁRIA NA COMUNIDADE CRISTÃ.

(2) No zoológico, tratadores escondem a comida pela jaula, PARA O ANIMAL PASSAR MAIS TEMPO PROCURANDO POR ELA, como na natureza.

Estes exemplos são do *corpus* da língua escrita. A oração em destaque em (1) é adverbial final desenvolvida por ser introduzida pela locução conjuntiva *para que* e por apresentar verbo no modo subjuntivo – *fossem*. Outro aspecto importante é o fato de o sujeito da oração final – *povos não-judaicos* – estar explícito e não ter correferência com o da principal – *ele*. Além disso, em relação à posição que ocupa, a oração final encontra-se posposta à principal. Esses aspectos não serão aprofundados agora, pois serão mais bem analisados no capítulo 5.

O exemplo 2 é uma oração final reduzida introduzida pela preposição *para*, cujo verbo assume a forma infinitiva – *passar* –, e que tem características estruturais idênticas às do exemplo (1): o sujeito explícito – *o animal* – é diferente (não correferente) do da principal, que é *tratadores*, e a oração final está também posposta à principal.

Outro aspecto importante que também será controlado nas análises aqui feitas é o traço sintático-semântico do sujeito da oração principal, pois num período composto em que se expressa a noção de finalidade, em termos de seu protótipo, espera-se que haja um sujeito que possa exercer algum controle sobre o que se projeta na oração final, conforme se viu em Neves (2000, p.888) no início deste capítulo. Desse modo, os exemplos (1) e (2) trazem mais esse aspecto em comum, já que os sujeitos da oração principal – *ele* e *tratadores* – atuam de maneira a desencadear os fatos expressos na oração final.

Frise-se que, com esse emparelhamento, o objetivo é demonstrar que uma oração adverbial final com determinadas características sintático-semânticas pode ocorrer na língua tanto na

forma desenvolvida quanto na reduzida. A frequência com que isso ocorre e os fatores que influenciam no uso de uma ou de outra forma serão também discutidos no capítulo 5.

Como se viu, a opção aqui foi de trabalhar com pares de exemplos reais. Contudo, numa tentativa intuitiva e prática, é possível reduzir o exemplo (1) e desenvolver o (2), cujas paráfrases são propostas abaixo:

(1)' [...] ele criou condições PARA POVOS NÃO-JUDAICOS, ao receberem a mensagem de Deus, SEREM INSERIDOS DE FORMA IGUALITÁRIA NA COMUNIDADE CRISTÃ.

(2)' No zoológico, tratadores escondem a comida pela jaula, PARA QUE O ANIMAL PASSE MAIS TEMPO PROCURANDO POR ELA, como na natureza.

Nas duas paráfrases, estão sublinhadas as mudanças feitas para que a redução de uma e o desenvolvimento da outra fossem realizados. Isso quer dizer que essas ocorrências têm potencial para ocorrer na língua de forma variável, num mesmo contexto e com o mesmo valor de verdade, sem que se altere seu significado referencial.

Os próximos dois exemplos do *corpus* da fala partilham das mesmas propriedades dos anteriores, e estão transcritos para que se perceba a possibilidade de ocorrerem nas duas modalidades da língua – a fala e a escrita –, de acordo com o gênero textual do qual os dados da língua em uso emergem.

(3) [...] eu tenho que gerenciar e organizar... fazer um esquema... PRA QUE TUDO FUNCIONE MUITO BEM [...]

(4) [...] tinha que voltar a tomar de novo daqui a seis horas, PRA FEBRE PASSAR NOVAMENTE [...]

Em síntese, a final desenvolvida em (3) e a reduzida em (4) estão constituídas de sujeitos explícitos não correferenciais; os sujeitos da oração principal têm o traço semântico “controlador”; ambas as orações finais estão dispostas na ordem canônica ou prototípica, ou seja, pospostas à principal. Além disso, parafraseá-las numa perspectiva variável – transformar a desenvolvida em reduzida e vice-versa – também é possível, desde que sejam feitas as devidas adaptações, como se vê a seguir:

(3)' [...] eu tenho que gerenciar e organizar... fazer um esquema... PRA TUDO FUNCIONAR
MUITO BEM [...]

(4)' [...] tinha que voltar a tomar de novo daqui a seis horas, PRA QUE A FEBRE PASSASSE
NOVAMENTE [...]

Na sequência, de forma mais concisa, serão apresentados mais alguns pares mínimos pertencentes aos dois *corpora*, os quais contêm configurações estruturais semelhantes na ocorrência entre finais desenvolvidas e reduzidas. Assim, em

(5) [...] o ensino público ele deve ser melhorado desde a base PRA QUE SEJA TÃO BOM
QUANTO AS ESCOLAS PARTICULARES [...]

e

(6) [...] pra gente de dentro do bairro não tem, ou, então, tem que sair de Consolação, ir
pra Praia do Canto PRA VER OS ESPORTES [...]

o que se tem são exemplos da base de dados da fala, em que (5) traz uma final desenvolvida e (6) uma reduzida. Seguindo a mesma linha da descrição feita anteriormente, as propriedades nelas identificadas são: sujeito da oração final não expresso e correferente ao da principal – *o ensino público* em (5) e *a gente* em (6), este recuperado anaforicamente; ordem canônica, isto é, oração final posposta à principal; e uma diferença em relação ao traço semântico do sujeito, que em (5) é paciente, enquanto em (6) é agente e controlador.

Uma paráfrase de (5) sob a forma da variante final reduzida apresenta um potencial considerável de ocorrer na língua em uso. Todavia, o mesmo não se pode afirmar em relação ao exemplo (6), pois a variante final desenvolvida com sujeito não expresso e correferente é rara. Essa questão está vinculada a um maior grau de integração entre as orações, o que é promovido, sobremaneira, pela correferencialidade do sujeito, algo que não está centralmente no escopo desta pesquisa, mas que já foi investigado em vários trabalhos, em diferentes línguas, como os de Haiman e Thompson (1984), Lehmann (1988), Braga (1996), Paiva (1999), Görski (2000), entre outros.

Seguindo com mais algumas ilustrações sobre as possibilidades de ocorrência das orações finais, examine-se outro par de dados da língua escrita:

(7) O rei concedia terra aos melhores guerreiros PARA QUE PUDESSEM SE SUSTENTAR e, em troca, os cavaleiros estavam sempre preparados para lutar em nome do soberano.

(8) Que ele mantém dois manequins vestidos de guarda à porta de seu quarto PARA IMPEDIR A ENTRADA DE FANTASMAS.

Nas ocorrências em (7), final desenvolvida, e em (8), final reduzida, uma configuração relativamente diferente do sujeito da oração final é identificada, pois não há plena correferencialidade, mas sim parcial. Isso significa que o sujeito da oração final não tem identidade com o sujeito da outra oração, mas retoma anafórica ou cataforicamente – dependendo da posição em que a final estiver disposta – algum termo que esteja exercendo outra função sintática na oração à qual a final estiver vinculada, daí a sua correferencialidade parcial. Em (7), o objeto indireto *os melhores guerreiros* é retomado e funciona como sujeito da oração final; já em (8), o objeto direto *dois manequins* passa a ser o sujeito na final reduzida. Portanto, as propriedades estruturais partilhadas neste par são: sujeito não expresso parcialmente correferente; sujeito da oração vinculada à final com traço agentivo/controlador; e orações finais em posição prototípica, ou seja, posposta.

Quanto às suas paráfrases, o fato de não haver correferência total do sujeito contribui para que elas sejam viáveis na língua em uso, de maneira que em (7) é possível projetar a redução oracional – *para poderem se sustentar* –, e em (8) o desenvolvimento – *para que impeçam a entrada de fantasmas*. No entanto, é importante assinalar que essas projeções, no uso efetivo da língua, poderiam sofrer determinados cortes ou mesmo inserções, de acordo com as necessidades do contexto discursivo. Em relação a esses casos, em especial, já que se trata de dados de um gênero da escrita revisada/monitorada, que se preocupa com a correção normativa, as suas paráfrases poderiam requerer, para garantir clareza ao texto, a expressão do sujeito. Assim, as paráfrases que estão sendo propostas têm apenas a finalidade metodológica de apontar potencialidades e não verdades absolutas e irretocáveis.

Como se verá na apresentação dos resultados, o número de finais desenvolvidas encontrado no PORTVIX foi pequeno, o que torna difícil estabelecer pares mínimos a partir dos exemplos reais dessa base de dados. Como a variabilidade na escrita se mostrou mais ampla,

segue a descrição do fenômeno em análise com mais um par de dados extraídos das reportagens:

(9) [...] PARA QUE O GELO SE FORME, é preciso que exista um primeiro cristal a partir do qual o resto da água congela.

(10) PARA COMPRAR UMA ARMA, é preciso apresentar "razões válidas", como ser membro de um clube de tiro.

Esses dois exemplos não constituem um par mínimo pleno, porque apresentam uma propriedade divergente, que é o sujeito explícito na final desenvolvida em (9) e o não explícito na reduzida em (10). No entanto, foi o mais próximo que se conseguiu chegar para que sejam verificadas duas configurações ainda não descritas ao mesmo tempo. Como se pode perceber, a estrutura das orações às quais a final está vinculada apresenta um tipo de sujeito que não traz as características prototípicas de quando se emprega um período composto que contenha oração final, já que esse sujeito é oracional nos dois exemplos. Além disso, pela primeira vez, estão sendo expostos dados em que a oração final está colocada antes da principal. A anteposição, como será visto no capítulo 5, foi bem menos frequente que a posposição nos dados levantados, mas nem por isso a posição da oração final deixou de ser levada em consideração na análise dos dados.

É importante pontuar que os 10 exemplos até aqui expostos e descritos não constituem as únicas configurações encontradas, nas bases de dados consultadas, para as orações finais desenvolvidas e reduzidas. Outras associações entre as propriedades sintático-semânticas que foram consideradas de importante controle e observação também aconteceram e serão apresentadas oportunamente no momento da análise dos resultados.

A partir deste ponto, depois de uma exposição mais detalhada do fenômeno linguístico que é o foco deste trabalho, serão apresentados os objetivos que o norteiam, bem como as ideias que motivaram o seu desenvolvimento e as principais hipóteses levantadas antes da análise propriamente dita dos dados e das conclusões aferidas.

1.2 Por quê? Para quê? O que se projeta?

Num primeiro momento, como sucede na maioria dos casos, a proposta de desenvolvimento desta pesquisa foi motivada pela observação dos fatos linguísticos por parte do pesquisador.

Particularmente no que tange ao fenômeno da alternância entre as finais desenvolvidas e reduzidas, houve outro fenômeno linguístico que antecedeu esse recorte: o fato de a língua portuguesa propiciar a seu usuário a possibilidade de empregar formas verbais no infinitivo pessoal flexionado, algo raro entre as línguas neolatinas. Mais especificamente, como mostra o seguinte dado da língua falada:

[...] então por isso que eu preciso que vocês me apoiem PARA COMPRARMOS MAIS BRINQUEDOS ... MAIS JOGOS ... MAIS BALANÇOS [...]

Neste exemplo, o verbo *comprar* foi empregado com a anexação da desinência número-pessoal *-mos*, o que lhe confere o *status* de infinitivo pessoal flexionado⁵. Esse aspecto suscitava mais curiosidade ainda quando se pensava em línguas como o espanhol, o francês e o italiano, que têm uma morfossintaxe semelhante à do português em certos pontos, mas não apresentam tal recurso de flexão em seu sistema – pelo menos, não na língua em uso⁶. Assim, ainda em primeira instância, os olhos se voltavam justamente para aquilo que soava como

⁵ Essa maneira de flexionar o infinitivo muitas vezes é confundida com o tempo *futuro do subjuntivo*, principalmente em verbos regulares, como é o caso de *comprar*. Entretanto, essa dúvida pode ser minimamente solucionada se se fizer um teste aleatório com qualquer verbo irregular, que se verá a diferença entre o infinitivo flexionado e o futuro do subjuntivo. Tomando-se de empréstimo o exemplo acima e substituindo-se “para comprarmos” – que é forma tanto do infinitivo flexionado quanto do futuro do subjuntivo, diferenciando-se apenas quando inseridas num contexto discursivo – por “para fazer”, que é um verbo irregular, o que se obtém é “para fazermos”, forma do infinitivo flexionado, portanto. Ao passo que, se o verbo “fazer” estivesse no futuro do subjuntivo, seria obtida a forma “fizermos”, impossível de ocorrência neste contexto.

⁶ No exemplo da fala do português, vê-se que o interlocutor se junta às outras pessoas para comprar os brinquedos, o que fica claro pela flexão do verbo comprar em “comprarmos”. Numa tradução livre para o espanhol desse mesmo trecho, mantendo o mesmo viés discursivo, o resultado seria algo como: “[...] *es por eso que necesito que me apoyen para que compremos más juguetes...*”. A forma “comprarmos” inexistente no espanhol em uso nos diferentes países que têm essa língua como idioma oficial. Para se expressar em contextos como o descrito, o usuário da língua teria, portanto, de recorrer categoricamente a uma oração final desenvolvida com o verbo no presente do subjuntivo.

diverso do comum: essas formas flexionadas do infinitivo, que podem emergir discursivamente nas orações reduzidas. Além disso, parecia interessante aos ouvidos o fato de que, em muitas ocorrências da língua em uso, o mesmo infinitivo pudesse ser empregado com ou sem flexão no mesmo contexto, notoriamente na primeira pessoa do plural, como no exemplo anterior, ou na terceira do plural, como em

Só que os físicos, cosmologistas e astrônomos não são pagos PARA TEREM BOM SENSO – sua obrigação é investigar o cosmo com todo rigor e descobrir do que ele é realmente feito, por mais estranho que possa parecer.

Neste exemplo da escrita, é possível vislumbrar a ocorrência de *ter* no lugar de *terem*. Ou seja, via-se aí um comportamento linguístico variável. Entretanto, em função de certa dificuldade em encontrar um número razoável de dados para o trabalho – como aponta Labov (2008, p. 239) ao falar do problema da “[...] raridade das formas gramaticais cruciais necessárias para a composição dos dados [...]” –, o que de início despertou a curiosidade levou a um conjunto mais amplo de elementos para compor o quadro desta pesquisa.

Que elementos? Um deles foi o tipo oração em que se pensou para traçar as estratégias de busca e seleção dos dados, que nesse caso foram as adverbiais finais, já que são orações que se reduzem apenas de infinitivo, o que facilitaria o recorte a ser trabalhado. No entanto, viu-se na configuração da pesquisa que a grande questão não residia no infinitivo em si – tanto que a sua flexão nem foi controlada neste trabalho –, mas em toda a composição da estrutura linguística das orações finais desenvolvidas e reduzidas, de maneira que o fenômeno a ser descrito, em seu comportamento variável, sintático e funcional, passou a ser o desenvolvimento e a redução de orações subordinadas adverbiais finais, reconhecendo-se nelas os elementos gerais e específicos que atuam para a ocorrência mais ou menos frequente de uma ou de outra forma, e em que domínios discursivos elas emergem.

É interessante pontuar que nesse trajeto entre a primeira ideia de pesquisa até se chegar à definitiva, um aspecto despertou a atenção: com base no princípio linguístico da marcação (GIVÓN, 1995), que se verá no capítulo 4 com mais detalhes, intuitivamente se identificava a final desenvolvida como o caso marcado, pois esse tipo de dado, quando encontrado, soava como diferente em razão de sua menor frequência. Esse já seria um motivo para talvez não considerar o fenômeno como passível de tratamento variacionista; contudo, como a proposta

sempre foi trabalhar com duas bases de dados, uma da fala e outra da escrita, em termos comparativos, a empreitada foi levada adiante, mesmo que na fala a área de variação em princípio fosse menor e na escrita, maior – ao menos nos gêneros textuais dos *corpora* selecionados.

Nesse percurso, como se trata de uma pesquisa de análise e descrição de fatos linguísticos, importa significativamente a observação dos dados da língua “flagrados” em seu funcionamento. Assim, por se ajustar à interpretação do fenômeno linguístico deste estudo, é justificável a utilização de alguns princípios teóricos do funcionalismo norte-americano, que

[...] concebe a linguagem como instrumento de interação social e [...] seu interesse de investigação linguística vai além da estrutura gramatical, buscando no contexto discursivo a motivação para os fatos da língua. A abordagem funcionalista procura explicar as regularidades observadas no uso interativo da língua analisando as condições discursivas em que se verifica esse uso (FURTADO DA CUNHA; COSTA; CEZARIO, 2003, p. 29).

É justamente nesse *contexto discursivo* e *uso interativo da língua* que se projetou perceber a dinâmica da alternância entre as adverbiais finais desenvolvidas e reduzidas. Para tanto, a maneira escolhida para se chegar a uma percepção mais concreta desse fenômeno foi o tratamento quantitativo dos dados linguísticos, feito mediante a teoria da sociolinguística variacionista. Essa atitude metodológica tem seu fundamento na medida em que se entende que esta pesquisa está trabalhando com um fenômeno de natureza variável. Assim,

A variação linguística, entendida como alternância entre dois ou mais elementos linguísticos, por sua própria natureza, não pode ser adequadamente descrita e analisada em termos categóricos ou estritamente qualitativos. [...] O uso de métodos estatísticos [...] tem permitido demonstrar o quão central a variação pode ser para o entendimento de questões como identidade, solidariedade ao grupo local, comunidade de fala, prestígio e estigma, entre tantas outras. (GUY; ZILLES, 2007, p. 73)

É importante frisar, contudo, que a análise do fenômeno de variação das orações finais, feita neste trabalho, não busca identificar questões relacionadas à identidade, ao prestígio e ao estigma linguístico, porque o próprio fenômeno está sendo visto como mais amplo no português e como governado muito mais por aspectos sintáticos e cognitivos que propriamente por aspectos sociais.

Como não poderia ser diferente, além das motivações individuais e teóricas, este trabalho tem a intenção de contribuir para os estudos na área da sintaxe da língua portuguesa, em particular do português brasileiro, com a pretensão de apontar caminhos para o entendimento de um fenômeno sintático de caráter variável, não apenas na oposição final desenvolvida x reduzida, mas também na composição dos elementos que perfazem a sua estrutura. Obviamente, não se trata de uma pesquisa totalmente inédita acerca desse fenômeno, como será visto no capítulo 3, em que se falará sobre alguns trabalhos que elegeram as orações finais como objeto de estudo; contudo, o diferencial que se propôs a fazer ao longo de seu desenvolvimento foi aprofundar a análise dos fatores que determinam a área de variação delimitada, interpretando-os à luz da probabilidade estatística e das formas linguísticas advindas do contexto discursivo, e projetá-la na comparação entre gêneros textuais diferentes de modalidades também diferentes da língua, a fala e a escrita. Em relação à fala, pela primeira vez são analisados dados dessa natureza do *corpus* PORTVIX, o qual contém dados referentes a entrevistas de informantes da cidade de Vitória-ES, gravadas segundo a metodologia laboviana.

Desse modo, para que este estudo se desenvolvesse, elaborou-se o seguinte objetivo geral para norteá-lo:

- observar e descrever como se organizam e são empregadas, em termos de sua variação sintática, as orações subordinadas adverbiais finais desenvolvidas e reduzidas, em dados da fala e da escrita do português brasileiro, por meio da abordagem teórica da sociolinguística variacionista e de alguns princípios do funcionalismo norte-americano.

Para que pudesse ser alcançado, esse objetivo geral foi desdobrado nos seguintes específicos:

- identificar, por meio da quantificação dos dados, a maior ou menor frequência das variantes da variável linguística em análise, de acordo com os gêneros textuais que compõem os *corpora* dos quais os dados da pesquisa foram extraídos;
- verificar se e como a variação sintática do fenômeno é motivada simultaneamente por variáveis internas e externas ao sistema linguístico;
- analisar qualitativamente as variáveis linguísticas estatisticamente relevantes ou não, identificadas por meio do tratamento quantitativo dos dados, devido à sua importância para o entendimento do objeto de estudo;

- comparar os resultados encontrados na fala e na escrita, para que sejam averiguados os pontos de convergência e divergência nas ocorrências das orações finais numa e em outra modalidade da língua;
- apontar aspectos funcionais do uso da língua que governam o fenômeno do desenvolvimento e da redução das orações finais;
- entender o fenômeno linguístico como não discreto (finais desenvolvidas x finais reduzidas), mas atribuir-lhe um tratamento escalar, com base no princípio da marcação linguística.

Conhecidas as principais razões que levaram ao desencadeamento deste trabalho, bem como os objetivos que o perpassam, vêm a seguir as principais hipóteses projetadas a respeito do fenômeno em análise, motivadas e encaminhadas em função dos primeiros contatos com os dados.

Com base numa primeira observação dos dados e em leituras prévias de outros trabalhos, a expectativa é de que as finais reduzidas sejam as mais frequentes na língua em uso, tanto na fala quanto na escrita. Ainda assim, espera-se encontrar um número maior de desenvolvidas na escrita que na fala, em razão do gênero textual escolhido – reportagem de cunho científico –, em que há uma elaboração mais formal da língua. A ideia é de que, com a predominância de uma linguagem mais impessoal e formal, as finais desenvolvidas possam emergir no discurso com um pouco mais de frequência.

Como já se frisou, um dos pontos centrais deste trabalho é verificar que fatores exercem maior influência sobre a área de variação do fenômeno em análise. Por uma percepção preliminar, aspectos relativos ao sujeito da oração final têm papel importante na configuração variável das desenvolvidas e reduzidas, especificamente no tocante à sua (não) correferencialidade com o sujeito da oração principal e à sua (não) explicitação. Isso porque, de acordo com o princípio da economia linguística (HAIMAN, 1983), segundo o qual se codifica menos o que se deduz pelo contexto discursivo, se o sujeito da oração final for correferente ao da principal, a tendência seria a da sua não explicitação, configuração esta que, em tese, tenderia a favorecer o uso de finais reduzidas. Essa tendência seria, então, motivada por iconicidade (HAIMAN, 1983), já que uma menor codificação na oração final faria com que esta apresentasse uma estrutura menor, ou seja, reduzida, sem a locução conjuntiva e com verbo no infinitivo. Nesse sentido, pelas mesmas motivações, o esperado é que a ocorrência de sujeito não correferencial favoreça a sua explicitação e, por conseguinte,

o uso da final desenvolvida. A questão é ver como essa associação se dará nos dados dos *corpora* investigados, bem como observar se esses aspectos relativos ao sujeito da oração final podem se distribuir de outras maneiras que não as apresentadas.

Em relação à posição em que se encontra a oração final, se anteposta ou posposta à principal, a sua variabilidade deve depender bastante do contexto discursivo dos quais os dados foram extraídos, pois essa questão em muito está vinculada às diferentes modalidades da língua, fala e escrita, representadas neste trabalho por gêneros textuais também diferentes. O que se espera, de forma ampla, é que a posição posposta seja a mais frequente, já que canonicamente a finalidade vem expressa após o que se declara, na oração principal, para que ela seja alcançada. A expectativa é, portanto, averiguar se na fala ou na escrita, ou em ambas, a anteposição ou a posposição terão algum efeito sobre o desenvolvimento e a redução das orações finais.

Outra observação que não se configura propriamente como uma hipótese, mas como um critério importante para a deflagração de uma hipótese, é o entendimento de que o fenômeno de variação sintática aqui pesquisado não é discreto, ou seja, que não há uma configuração única para uma oração desenvolvida e outra para uma reduzida. O que se preconiza, assim, é que há uma escala entre elas. Essa percepção é decorrente, portanto, da consideração também escalar do princípio da marcação (GIVÓN, 1995) – que será discutido no capítulo 3. Após a apuração quantitativa dos dados, será utilizado o critério da frequência das ocorrências, que está embutido no princípio da marcação, para que sejam aferidas algumas escalas, tomando-se por base a associação de alguns dos grupos de fatores utilizados no tratamento quantitativo e estatístico dos dados (cf. seção 5.5).

No que diz respeito a dois dos fatores sociais selecionados, a idade e o sexo/gênero do usuário da língua, supõe-se que eles pouco tenham a ver com a alternância entre as variantes da variável linguística em estudo. No entanto, há alguma expectativa em relação à terceira variável social, que é o nível de escolarização dos informantes, pois os entrevistados – amostra da fala – com nível universitário de escolaridade, e os jornalistas – amostra da escrita –, todos com curso superior, de início parecem fazer mais uso de orações desenvolvidas que usuários com menos tempo de escolarização, o que em tese pode vir a trazer algum efeito sobre a variação das orações finais. Isso porque se está entendendo que, pelo fato de as finais desenvolvidas serem mais complexas estruturalmente e requererem o emprego do verbo no modo subjuntivo, elas seriam mais típicas nas ocorrências daqueles que tenham tido mais

acesso a leituras de textos de natureza diversa. Todavia, não se espera, mesmo nesse nível de escolarização, que as desenvolvidas superem a frequência de uso das reduzidas.

Por outro lado, é necessário dizer, não é pretensão deste trabalho vincular a ideia de marcas de identidade numa comunidade de fala ao objeto linguístico deste estudo – atitude bastante frequente em pesquisas na área da sociolinguística variacionista. Tampouco se acredita que a alternância entre o uso de uma oração desenvolvida ou reduzida possa trazer certo grau de estigma social para este ou aquele usuário, conforme acontece com muitos outros fenômenos linguísticos, como no não estabelecimento das relações de número na concordância verbal e nominal, ou no uso muito frequente do gerúndio – neste último caso, várias discussões inflamadas em âmbito nacional têm sido levantadas, por especialistas ou não.

Ao falar da importância da variável escolaridade, Votre (2008, p. 52) diz que

A observação do dia a dia confirma que a escola gera mudanças na fala e na escrita das pessoas que as frequentam e das comunidades discursivas. Constata-se, por outro lado, que ela atua como preservadora de formas de prestígio, face a tendências de mudança em curso nessas comunidades. Veículo de familiarização com a literatura nacional, a escola incute gostos, normas, padrões estéticos e morais em face da conformidade de dizer e de escrever.

Ainda nessa visão social aplicada à percepção que se tem da língua em uso, um aspecto chama a atenção: usar mais ou menos reduzidas, mais ou menos desenvolvidas pode não gerar constrangimentos sociais; mas, ao que parece, o emprego destacado de desenvolvidas por um falante – quando a aplicação do subjuntivo é obrigatória⁷ –, ou ainda o uso das reduzidas com infinitivo pessoal com marca de flexão, como acontece na primeira e na terceira pessoas do plural (*dizermos* e *dizerem*, por exemplo) podem, ao contrário, ser indicados empiricamente como formas de prestígio social.

Como afirma Votre (2008), talvez nesse ponto é que o nível de escolarização possa ser determinante para que o usuário da língua com maior tempo de frequência à escola empregue mais as duas estruturas linguísticas mencionadas – as desenvolvidas e as reduzidas com flexão do infinitivo. Entretanto, buscar essa constatação a respeito do infinitivo flexionado não faz

⁷ Neste caso, fala-se em subjuntivo de maneira generalizada, porque as finais podem apresentar o verbo tanto no presente quanto no pretérito imperfeito do subjuntivo.

parte do escopo desta pesquisa, surgindo aqui apenas como forma de ponderação sobre a possível influência do nível de ensino nos usos linguísticos.

No próximo capítulo será feita uma exposição a respeito da maneira como são apresentadas e tratadas as orações finais nas gramáticas de caráter normativo e descritivo.

CAPÍTULO 2

AS ORAÇÕES FINAIS NAS GRAMÁTICAS NORMATIVAS E DESCRITIVAS

Em geral, nos capítulos dedicados ao estudo da sintaxe, particularmente na parte relativa aos períodos simples e compostos, é que as gramáticas normativas relatam algo a respeito das orações subordinadas desenvolvidas e reduzidas. Na verdade, durante toda a exposição do período composto (por coordenação e por subordinação), é de praxe que o gramático trabalhe apenas com exemplos de orações desenvolvidas, para, no final do capítulo, voltar-se para as “chamadas orações reduzidas” (BECHARA, 2009, p. 513). Não é muito lembrar que, como rege a metodologia de trabalho da tradição gramatical, em grande parte dos casos só se encontram exemplos da intuição do gramático e fragmentos de textos de autores canônicos das literaturas portuguesa e brasileira, além de não haver preocupações de cunho discursivo e contextual das ocorrências. Nesse sentido, serão vistas, a seguir, algumas considerações encontradas em algumas dessas gramáticas a respeito das noções de oração subordinada desenvolvida e oração reduzida, e, em particular, sobre o tipo de oração selecionada para este estudo: as adverbiais finais.

Em Bechara (2009), há um destaque incomum para as orações reduzidas em relação ao que se encontra em outros manuais com fins normativos. O autor define oração desenvolvida como aquela em que a(s) forma(s) verbal(is) se encontra(m) em modos e tempos distintos das formas nominais dos verbos da língua portuguesa. Por outro lado, chama de reduzida a oração que tenha “[...] verbo (principal ou auxiliar, este último nas locuções verbais) [...] no infinitivo, gerúndio e particípio (reduzidas infinitivas, gerundiais e participiais)” (BECHARA, 2009, p. 513). Para o gramático, pode-se considerar reduzida aquela oração que contenha uma das formas nominais mencionadas, que tenha autonomia sintática no enunciado em que esteja inserida e, ainda, que apresente a possibilidade de ser estruturada tal como uma oração desenvolvida, ou seja, com verbo em alguma forma finita. Para melhor compreender as

colocações de Bechara, veja-se o seguinte exemplo⁸ de uma final reduzida (1) e um possível desdobramento seu em final desenvolvida (2):

(1) Ela para de agir assim que o intestino está liberado, abandonando você à própria sorte
PARA RESISTIR ÀS TENTACÕES.

(2) Ela para de agir assim que o intestino está liberado, abandonando você à própria sorte
PARA QUE (VOCÊ) RESISTA ÀS TENTACÕES.

A respeito da questão do desdobramento das reduzidas em desenvolvidas, Bechara (2009), com a voz na estilística, afirma que esse recurso permite àquele que redige em português expressar-se em estilos variados, de maneira a apresentar uma escrita mais concisa e sem acumulações de “quês” – isso no caso da redução, em que necessariamente o conectivo “que” é retirado da construção.

Entretanto, o autor problematiza algo que não é consenso entre gramáticos e linguistas: o conceito de oração reduzida. Segundo ele, muitos partem “[...] da ideia de que o que caracteriza a relação prediativa é a presença de verbo na forma finita [...]” e, assim, “[...] a construção com verbo nas formas nominais (infinitivo, gerúndio e particípio) não constitui oração, e, neste caso, é uma subunidade da oração, um termo dela, quase sempre como um adjunto adverbial” (BECHARA, 2009, p. 513). Apesar dessa colocação, o autor mantém a classificação e a nomenclatura de largo uso e, também, o *status* oracional que as reduzidas têm.

Melo (1968) é um dos gramáticos que discutem a natureza oracional das reduzidas. Para o autor, a propriedade nominal do infinitivo quer dizer que não há orações reduzidas de infinitivo. Assim, o expediente do desdobramento de reduzidas em desenvolvidas – anteriormente ilustrado em Bechara (2009) –, que tem por finalidade a constatação de que se trata do mesmo tipo de oração, não é aprovado pelo gramático, atitude chamada por ele de “processo de muletas”. Entretanto, parece contradizer-se quando usa o exemplo “Bento, ao pular-lhe o coração em saltos de ruim presságio, ainda deu três passos para chamar o filho” (MELO, 1968, p. 83). A contradição está no fato de dizer que o termo “o coração” é sujeito da

⁸ Como não se quis usar exemplos que não fossem do português brasileiro em uso, seja ele escrito ou falado, optou-se por selecionar, também para este momento em que se fala das gramáticas normativas, um exemplo do *corpus* da língua escrita que compõe esta pesquisa.

forma “pular”, classificação esta incoerente, já que o autor considera o infinitivo um nome e não um verbo, não se constituindo a estrutura, portanto, como oração.

Outra gramática normativa que será brevemente comentada é a de Cunha e Cintra (2001). Os autores não fazem extensas considerações sobre o fenômeno da redução de orações, sendo, ao contrário, breves e concisos. Dizem que é desenvolvida a oração que se inicia por um conectivo de natureza subordinativa e que apresenta verbo em algum tempo do modo indicativo ou do subjuntivo, e que é reduzida “[...] a oração dependente que não se inicia por relativo nem por conjunção subordinativa, e que tem o verbo numa das formas nominais – o infinitivo, o gerúndio, ou o particípio” (CUNHA; CINTRA, 2001, p. 609, 610). A seguir, expõem uma série de exemplos extraídos da literatura clássica, já que essa enumeração tem fins meramente normativos para regular a escrita formal. Vale dizer, ainda, que não fazem menção especial às orações finais, objeto deste estudo, apenas colocadas como um tipo de oração adverbial e exemplificadas.

Apesar de não constituir uma gramática normativa que traga as partes centrais de todo e qualquer manual desse tipo – fonética/fonologia, morfologia e sintaxe –, Kury (2008) aproxima-se bastante desse modelo, principalmente pelo tipo de dado com o qual trabalha: exemplos extraídos da literatura canônica. Basicamente, não apresenta diferenças nos conceitos de orações desenvolvidas e reduzidas relativamente aos dados por Bechara (2009) e Cunha e Cintra (2001). Mesmo assim, traz em sua composição apontamentos mais específicos relativos às orações reduzidas e às adverbiais finais. Sobre estas, Kury (2008, p. 99) diz que as finais desenvolvidas “[...] são introduzidas por uma conjunção ou locução conjuntiva final, cujo tipo é *para que* (*a fim de que, que, porque*)”, e que as finais reduzidas apresentam o verbo sempre no infinitivo e costumam ser iniciadas pelas preposições *para, a, de, por*, ou pela locução *a fim de*.⁹

Rocha Lima (2010, p. 325) faz em seu manual uma sucinta classificação e conceituação a respeito das orações subordinadas quanto à forma e ao modo como se articulam com a oração principal, a saber: desenvolvidas, reduzidas e justapostas. Apresenta também a possibilidade de as desenvolvidas poderem ser transformadas em reduzidas, bem como o contrário desse processo. Mas faz algumas restrições para alguns tipos oracionais, como as adverbiais de

⁹ É preciso destacar que todos os manuais mencionados, ao apontarem os principais conectores que introduzem as orações finais (seja desenvolvidas ou reduzidas), limitam-se à escrita formal, literária e revisada. Como se verá mais adiante, os dados desta pesquisa mostram que tanto no *corpus* da oralidade quanto no da escrita a predominância das ocorrências são dos conectores *para* (*pra*) e *para que* (*pra que*).

modo, que, segundo ele, só ocorrem sob a forma reduzida. Em relação às finais, o autor apenas indica os principais conectivos que as introduzem, bem como aponta a possibilidade de elas ocorrerem sob as formas desenvolvida e reduzida.

No tocante às gramáticas descritivas, serão exploradas algumas das que trazem aspectos relevantes, com preocupações de caráter linguístico importantes para o desenvolvimento desta pesquisa. A primeira delas é a obra de Neves (2000), que foi elaborada com o objetivo de trazer para o campo dos estudos linguísticos uma amostra significativa de aspectos gramaticais gerais do português brasileiro no que diz respeito aos usos da língua escrita. O *corpus* em questão se baseia em textos de natureza romanesca, jornalística, técnica e também dramática e da oratória, que compõem um banco de dados de 70 milhões de ocorrências, armazenado no Centro de Estudos Lexicográficos da Universidade Estadual Paulista (Unesp – Araraquara). Embora os textos dramáticos se aproximem das características da oralidade, não constituem, *a priori*, um banco de dados representativo para o estudo sobre variação e mudança da fala. Há casos, no entanto, em que pesquisadores lançam mão desse recurso, com as devidas ressalvas, quando da impossibilidade de elaboração ou acesso a um *corpus* oral.

Inspirada em pressupostos teóricos funcionalistas, Neves (2000), em sua obra, busca apresentar um caráter descritivo da língua escrita em uso, mesmo seguindo a divisão e a classificação que normalmente aparecem nas gramáticas prescritivas tradicionais, algo que não desabona seu caráter de importância, pois essa escolha tem todo um sentido: facilitar a consulta tanto dos especialistas quanto daqueles que são de outras áreas de estudos e que queiram dela fazer uso. Neves (2000) insere, ao longo da obra, vários excertos desse banco de dados a fim de poder exemplificar determinada descrição que se faça na abrangência das duas grandes áreas de maior destaque na obra: a morfologia e a sintaxe.

Na obra, a autora organiza uma seção sob o rótulo “as construções finais”, em que estão incluídas as adverbiais finais, para as quais se chama a atenção neste trabalho. Assim, na sequência, serão expostas algumas considerações mais concretas de Neves (2000) a esse respeito, pois ela faz uma detalhada descrição a respeito dessas construções, sobretudo no que diz respeito às formas verbais nelas empregadas. Para tanto, foram selecionados dois aspectos descritivos para entender o que a autora propõe:

- (1) As orações finais desenvolvidas, normalmente iniciadas pela locução conjuntiva *para que*;

(2) As orações finais reduzidas, em geral iniciadas pela preposição *para*.

De acordo com Neves (2000, p. 886), o caso em (a) apresenta o verbo da oração subordinada em modo finito, e o seu sujeito é diferente do que ocorre na oração principal. Para que isso se verifique, segue o exemplo arrolado pela autora:

(a) **A leitura** serve **PARA QUE A CRIANÇA** APRENDA RAPIDAMENTE O QUE NÃO PODERIA ALCANÇAR SÓ.¹⁰

Vê-se que a oração principal tem como sujeito o sintagma *a leitura*, enquanto na final o sujeito é *a criança*. Assim, o verbo da final é empregado em modo finito (presente do subjuntivo) pelo fato de ter sido introduzida por *para que*.

Para o caso em (2), Neves (2000, p. 887) afirma que “[...] as orações finais iniciadas por preposição (em infinitivo) não têm restrição quanto ao sujeito”. Assim, no *corpus* com o qual trabalhou a autora, ocorreram períodos compostos com sujeitos correferentes e não correferentes entre as orações que os compunham. Eis alguns exemplos coletados:

(b) **O prisioneiro** foi trazido para a praça **PARA SER LINCHADO**.

(c) **Convém**, atualmente, subdividir a própria contabilidade, **PARA MELHOR APREENDERMOS** A SUA FINALIDADE e **AVALIARMOS** A SUA EXTREMA IMPORTÂNCIA NA ADMINISTRAÇÃO MODERNA [...]

Em (b), verifica-se que o sujeito da oração principal e da final é o sintagma *o prisioneiro*, enquanto que em (c) há um sintagma oracional na função de sujeito da forma *convém*, enquanto nas construções seguintes as formas verbais fazem referência ao sujeito implícito *nós*. Nesse sentido, algo muito importante deve ser observado: a forma verbal empregada. Tanto em (b) quanto em (c), em função do uso da preposição *para*, o infinitivo foi utilizado. Em especial, em (c) a marca da primeira pessoa do plural teve de ser expressa para tornar o conteúdo comunicativo mais pessoal, uma vez que empregar as formas *aprender* e *avaliar* o deixaria com marcas de impessoalidade. Não se pode afirmar categoricamente – pois apenas

¹⁰ Todos os exemplos da parte referente a Neves (2000) são do banco de dados da língua escrita, já especificado anteriormente.

se tem o fragmento extraído por Neves –, mas parece haver aí um uso funcional da forma verbal para que a devida referência fosse efetivada, com base nos propósitos comunicativos do autor do texto. De acordo com o descrito acima, não houve ocorrências de casos em que foi empregada a locução conjuntiva *para que* quando os sujeitos de duas ou mais orações eram correferentes. Assim, parafraseando o exemplo de Neves (2000), a frase em (b) não apareceria sob a forma

(d) **O prisioneiro** foi trazido para a praça **PARA QUE FOSSE LINCHADO**.¹¹

Isso ocorre porque, segundo a descrição já apontada, com a locução *para que* os sujeitos teriam de ser necessariamente não correferentes: “As **orações finais** iniciadas por **conjunção** (em **modo finito**) constroem-se com **sujeito** diferente do da **oração principal**” (NEVES, 2000, p. 886). Essa definição é um tanto curiosa porque parece ser que o emprego de *para que* é que motiva a explicitação do sujeito diferente (não correferente), quando o que se pretende ver neste trabalho é justamente uma motivação inversa, ou seja, verificar os reais efeitos da explicitação ou não do sujeito sobre a ocorrência de uma das formas de realização das orações finais.

A autora ainda apresenta uma série de características e propriedades das adverbiais finais, tais como dizer que elas podem ser *clivadas* e *focalizadas*. Essa estratégia corresponde a antecipar a final para o início do enunciado e nela inserir partículas expletivas ou de realce. Um dos exemplos utilizados por Neves (2000, p. 890) é o seguinte:

(e) Esboçou um movimento **PARA QUE SEGUÍSSEMOS EM FRENTE**.

Clivada (com o emprego de *foi/que*) e focalizada (anteposta), a construção assim ficaria:

(f) **Foi PARA QUE SEGUÍSSEMOS EM FRENTE** **que** esboçou um movimento.

¹¹ Mais adiante, será visto que, mesmo sendo casos pouco numerosos, há ocorrências desse tipo, sobretudo na fala informal. E esse exemplo apresenta uma particularidade: o traço [-agentivo] do sujeito, já que a oração final está estruturada na voz passiva. Esse fator parece fazer alguma diferença para a ocorrência, porque a frase (d) soa como de uso corrente em português.

Essa propriedade é importante no sentido de que a natureza adverbial da oração se torna mais evidente, pelo fato de ela poder ser deslocada de sua posição inicial, sem mudanças do sentido original, caracterizando uma relação mais frouxa, sintaticamente menos integrada, que os elementos circunstanciais costumam apresentar em relação aos termos da oração que modificam.

Apontamento também feito é o da subcategorização das finais em função da modalidade que podem apresentar. Para tanto, Neves (2000) observa o verbo da oração principal, para definir os subtipos das finais em: factuais, hipotéticas e contrafactuais. Nesse sentido, se o verbo da matriz estiver no presente (g) ou nos pretéritos perfeito (h) ou imperfeito do indicativo, a subordinada final é factual. Eis os exemplos (NEVES, 2000, p. 889):

(g) Vê que alguém, em uma pedra mais alta, lhe **faz** sinais nervosos **PARA QUE** SAIA DALI.

(h) Mas a viagem era fretada e nada **foi feito** **PARA** APURAR O MISTÉRIO.

Por outro lado, se os verbos da oração principal estiverem no futuro do presente (i) ou do pretérito (j) ou ainda no imperativo (k), as finais são *hipotéticas*:

(i) Ele **será** muito frágil **PARA QUE** ALGUÉM O POSSA MATAR.

(j) Prometeu que **tomaria** as providências necessárias **PARA QUE** DONA LEONOR NÃO TORNASSE A SE INTROMETER NA MINHA VIDA PARTICULAR.

(k) E **preste bem atenção**, **PARA QUE** DEPOIS NÃO SE ASSUSTE.

E são consideradas contrafactuais aquelas que têm o verbo da principal no futuro do pretérito composto (l):

(l) Eu não sabia, que diacho, do que é que me acusavam nem o que **teria feito** **PARA** SER TRATADO ASSIM.

Outra gramática de descrição do português, mas não com a mesma inspiração teórica da de Neves (2000), é a de Perini (1995). O fato de os exemplos arrolados pelo autor serem dados da intuição e não da língua em uso traz uma diferença significativa em relação à exemplificação feita por Neves – mesmo que nela os dados sejam apenas da língua escrita. Assim, ao abordar as orações complexas, Perini refere-se às orações reduzidas, mas não faz menção ao tipo especificado neste estudo. Utiliza a noção de *predicado complexo* para solucionar o problema do linguista ao definir se, em casos como “Os juízes decidiram considerar ilegal o imposto”, há uma ou duas orações, ou seja, se “decidiram considerar” é uma locução verbal – em que um verbo é auxiliar do outro – ou se são dois verbos que apontam suas “exigências” em relação aos complementos que possam vir a ter: “O critério se baseia nos traços de transitividade, partindo do princípio de que cada oração tem um predicado e, portanto, um conjunto de traços de transitividade” (PERINI, 1995, p. 127). A conclusão a que se chega fora da obra, portanto, é de que no exemplo há uma estrutura de predicado complexo, já que os dois verbos são transitivos diretos, de maneira que a oração “considerar ilegal o imposto” está reduzida porque o sujeito é correferente ao da oração principal, sendo desnecessária a construção desenvolvida “Os juízes decidiram que consideram ilegal o imposto”. Pode-se perceber, pelo exemplo, que o autor está trabalhando com uma oração subordinada substantiva objetiva direta reduzida, e não com uma adverbial. Configura-se aí, portanto, outro tipo de relação sintática, que não está sendo investigado neste trabalho.

Em resumo, esta seção trouxe uma breve explanação sobre o que apresentam algumas gramáticas normativas e descritivas do português sobre o fenômeno das orações subordinadas adverbiais finais em suas modalidades desenvolvida e reduzida. O próximo capítulo abordará o referencial teórico da linguística que dá suporte a esta pesquisa, bem como trará algumas notas a respeito de outros trabalhos que também elegeram as orações finais como objeto de estudo ou como parte dele.

CAPÍTULO 3

FUNDAMENTAÇÃO TEÓRICA E REVISÃO BIBLIOGRÁFICA

3.1 Pressupostos teóricos da sociolinguística variacionista e do funcionalismo norte-americano

Este trabalho de pesquisa no âmbito da descrição de fatos e fenômenos linguísticos – neste caso, o desenvolvimento e a redução de orações subordinadas adverbiais finais – apoia-se, em primazia, na Teoria da Variação e da Mudança Linguística (WEINREICH; LABOV; HERZOG, 2006 e LABOV, 2008), também chamada Sociolinguística Variacionista. Como contribuição teórica complementar, são utilizados alguns princípios do Funcionalismo linguístico norte-americano, principalmente a partir de Haiman (1983) e Givón (1995, 2001).

A escolha da sociolinguística de orientação laboviana está pautada no fato de ser esta uma teoria que busca trabalhar essencialmente com dados do uso real e efetivo de uma dada língua, procurando descrevê-la e entendê-la por meio de um arcabouço metodológico quantitativo. Essa maneira de enxergar uma realidade linguística possibilita ao pesquisador desenvolver seu trabalho de análise buscando identificar fatores que apresentem maior ou menor influência sobre as ocorrências variáveis de determinado fenômeno linguístico. É uma visão teórica que busca perceber um mesmo nível de importância para influências de fatores linguísticos e sociais para o entendimento da variação linguística.

A Sociolinguística Variacionista busca sistematizar a heterogeneidade linguística, tratando-a como não aleatória. O sistema disponibiliza ao usuário formas linguísticas possíveis para a sua comunicação na comunidade de fala da qual faz parte, de maneira que a escolha ou seleção de uma variante pode estar relacionada à atuação de diferentes forças, internas e externas ao sistema. Nesse sentido, a Sociolinguística procura tornar paralelos, em termos de suas influências, o linguístico e o social. Assim, a variação é importante na organização da expressão linguística humana, já que é vista como própria e inerente aos sistemas linguísticos.

O funcionalismo norte-americano encontra seu espaço neste trabalho por ter como fundamento a investigação linguística a partir dos dados da língua em uso em seus diferentes domínios discursivos. Para Givón (2001, p.13), numa visão funcionalista, a gramática de uma língua

codifica, simultaneamente, tanto a semântica proposicional quanto a coerência do discurso (pragmática). Esse é, deveras, um dos fatos da gramática-como-código que causam mais perplexidade: embora esteja inteiramente localizada na oração, seu escopo funcional não trata primariamente da informação proposicional que se aninha, obscura, na oração em que reside. Pelo contrário, a gramática trata predominantemente das relações de coerência entre a proposição (oração) e seu contexto mais amplo no discurso.¹²

Nesse sentido, para uma abordagem interpretativa do fenômeno em questão, serão utilizados alguns dos princípios que compõem a corrente de estudos denominada funcionalismo de orientação norte-americana, ou mais recentemente chamada de linguística centrada no uso, entre os quais estão a *marcação*, a *economia* e a *iconicidade* – que serão abordados de forma mais adequada nas próximas seções deste capítulo.

Nessa perspectiva, pode-se dizer que este trabalho alinha-se ao que Neves (1999) chama de *sociofuncionalismo*. De acordo com a autora, muitos grupos de linguistas brasileiros têm se valido desse vínculo para desenvolver seus trabalhos, e cita o grupo PEUL (Programa de Estudos sobre o Uso da Língua) como “o maior e mais antigo desses grupos [...] que tem uma orientação variacionista dominante, movendo-se, pois, no que alguns membros do grupo definem como sociofuncionalismo. O projeto aborda a variação linguística sob o prisma da função discursiva das variantes” (NEVES, 1999, p. 75).

O que se busca explicar e dizer a respeito de fenômenos de variação e mudança nessas duas correntes tem bastantes pontos de interseção, tornando mais evidente a vinculação entre elas:

A sociolinguística e o funcionalismo compartilham, entre outras, as seguintes concepções teórico-metodológicas: a) entendem que a língua possui natureza heterogênea e variável; b) consideram fundamental que o

¹² Tradução livre a partir do seguinte fragmento do texto original: “Grammar codes, simultaneously, both propositional semantics and discourse coherence (pragmatics). This is indeed one of the most baffling facts about grammar-as-code: Although it is located wholly in the clause, its functional scope is not primarily about the propositional information couched in the clause in which it resides. Rather, grammar is predominantly about the coherence relations between the propositional (clause) and its wider discourse context”.

analista examine a língua utilizada em contextos reais; c) defendem que a língua está continuamente mudando; e d) valem-se de tratamento empírico com quantificação estatística como evidência para atestar fenômenos de variação e de mudança (GÖRSKI; TAVARES; FREITAG, 2008, p. 102).

As autoras dizem, ainda, que muitos estudos no âmbito do funcionalismo têm se preocupado com o possível influxo de fatores sociais sobre o uso da língua, tais como o nível de escolarização, a faixa etária, a classe social e o sexo do falante, aspectos centrais para os sociolinguistas variacionistas. Além disso, dão destaque para o fato de que os dois modelos têm voltado, sobremaneira, sua atenção para a frequência de uso das estruturas e formas linguísticas. E, para embasar o que dizem, trazem as ideias parafraseadas de Labov (1994), da sociolinguística, e de Bybee e Hopper (2001), do funcionalismo:

Para a sociolinguística, o uso depende do ambiente linguístico/e ou do contexto social, o que define a natureza do sistema linguístico como probabilística e pressupõe o emprego de técnicas quantitativas para a observação das regularidades – em termos de frequência de uso – que o governam [...]. Para o funcionalismo, as formas são unidades de processamento, armazenadas, acessadas e constantemente afetadas pela experiência, inclusive pela frequência de ocorrência, pois a representação cognitiva pode ser alterada pela exposição a repetidas instâncias de uso de uma construção inovadora (GÖRSKI; TAVARES; FREITAG, 2008, p. 102).

Essa articulação sociofuncionalista possibilita uma importante constatação a respeito dos estudos linguísticos que dela fizeram uso: de acordo com Naro (1998, p. 110), reportando-se a Labov, “[...] a variação no uso da língua é um fenômeno regular, sujeito a restrições de natureza estrutural, internas ao sistema linguístico, e social, externas a esse sistema”. Nesse âmbito, a variação não deve ser vista como um fenômeno linguístico aleatório, isto é, ela pode exercer algum papel num contexto discursivo mais amplo, nas diferentes situações comunicativas em maior ou menor grau. Assim, a funcionalidade na variação de determinada forma linguística se configura quando “[...] os falantes usam mais certas variantes em contextos em que o uso da variante em questão, e não outra variante, colabora para transmitir um determinado sentido em nível semântico ou discursivo” (NARO, 1998, p.118).

Nesse sentido, já que o funcionalismo vê a língua como um instrumento de interação social, é em momentos de interação linguística que a heterogeneidade e a dinamicidade da língua podem levar a situações ou contextos de variação e mudança. Para Labov (2008, p. 252-263), a variação é uma propriedade peculiar, própria, inerente ao sistema – algo que o autor coloca

como ponto instigante ao leitor, ao tratar da “simplificação dos grupos consonantais e do sufixo de tempo passado” no inglês norte-americano. De acordo com Scherre (1996, p.39-40),

[...] a variação não é aleatória, mas sim governada por restrições linguísticas e não linguísticas. [...] os fenômenos linguísticos variáveis, aqueles expressos por duas ou mais variantes, apresentam tendências regulares passíveis de serem descritas e explicadas por restrições de natureza linguística e não linguística.

Ao divulgar seus trabalhos no campo da linguística, e conferir aos fatores externos à língua um caráter de grande importância – mas não apenas a eles –, Labov (2008) traz algumas constatações fundamentais para os estudos que buscam entender certos fenômenos linguísticos no que diz respeito ao seu uso e emprego variável. Até então centrada em questões com foco predominante na *competência* linguística dos falantes, a ciência linguística passa a incorporar uma visão mais ampla, mais abrangente, com grande interesse na investigação da variação de fenômenos linguísticos, assim como passa a receber críticas importantes ao fato de o *desempenho* linguístico dos falantes ser praticamente desprezado como material válido em pesquisas para a descrição das línguas. É uma crítica bem direcionada aos primeiros postulados teóricos da gramática gerativa, de Chomsky.

Assim, muitas justificativas eram dadas para que a fala real e espontânea do dia a dia fosse excluída do campo dos estudos linguísticos, dentre as quais podem ser citadas a dificuldade de lidar com a variação linguística nela existente; a chamada agramaticalidade da fala; a dificuldade de trabalhar com materiais gravados – algo que nos dias atuais seria um argumento sem qualquer fundamento. Desse modo, grande parte das análises linguísticas era feita com base em exemplos construídos pela intuição do linguista, em confiança plena em sua capacidade de prever formas concernentes ao seu sistema linguístico – a privilegiada competência linguística.

A partir da sociolinguística variacionista, a atenção se volta fortemente para o desempenho (como fonte de dados) em detrimento da competência linguística do falante que integra uma comunidade de fala. Para Labov (2008, p. 260), interessado em entender como a variação de determinados fenômenos linguísticos aponta caminhos para a mudança linguística, “[...] no curso da evolução linguística, a mudança caminha para se completar, e regras variáveis se tornam invariantes. Quando isso acontece, há inevitavelmente alguma outra mudança estrutural que compensa a perda de informação envolvida”. Segundo ele, isso só é passível de

observação se o linguista voltar o seu olhar para os dados que realmente possam atestá-lo: a fala cotidiana.

Labov (2008, p.236), ao salientar a necessidade de melhor entender determinado fenômeno de investigação linguística, afirma que “temos de olhar para os dados da fala cotidiana o mais perto e diretamente possível, e caracterizar seu relacionamento com as teorias gramaticais do modo mais acurado que pudermos, corrigindo e adequando a teoria para que ela se ajuste ao objeto visado”.

Neste trabalho, portanto, optou-se por se fazer um contraste das ocorrências variáveis entre as orações finais desenvolvidas e reduzidas em dados coletados da fala espontânea (gênero entrevista – PORTVIX) e da escrita revisada (gênero reportagem – revista Superinteressante). Um dos principais objetivos dessa tarefa é verificar se há mais coincidências ou discrepâncias entre as restrições linguísticas e não linguísticas apresentadas por cada uma das modalidades da língua em seus respectivos gêneros. É um misto do que assevera Labov (2008) – ao falar da importância de flagrar o dado no vernáculo – com uma tendência cada vez mais forte nos estudos que se intitulam funcionalistas, que é averiguar as ocorrências linguísticas e entendê-las como tendo comportamentos específicos dentro de determinado gênero em cada modalidade da língua em uso, na fala ou na escrita.

Na próxima seção, são desenvolvidas algumas considerações a respeito dos princípios funcionalistas que servem de base para a interpretação do fenômeno variável adotado como objeto de pesquisa nesta dissertação.

3.2 Um viés funcionalista: os princípios da marcação, economia e iconicidade.

3.2.1 Marcação

Fazer referência à noção de marcação é uma atitude analítico-descritiva bastante frequente nos estudos linguísticos contemporâneos. Nesse sentido, a decisão de se considerar determinada forma ou estrutura linguística como mais ou menos marcada que outra é feita com base em alguns critérios: a *complexidade formal* dessa estrutura – incluindo-se aí

aspectos como a sua extensão ou mesmo o número de morfemas que contém –; a *frequência* com que ocorre; e a *complexidade cognitiva*, referente ao tempo de processamento para a elaboração da estrutura no momento do ato comunicativo, e também à demanda de maior atenção e maior esforço mental por parte do usuário da língua (GIVÓN, 1995, p. 28). Dessa maneira, as formas marcadas são aquelas mais complexas formalmente, menos frequentes nos dados da língua em análise e com tempo de processamento cognitivo mais extenso – deduz-se daí que o item não marcado é o que apresenta menor complexidade estrutural, é mais frequente em termos de seu uso e menos complexo do ponto de vista cognitivo.

Em relação a esses critérios que definem itens linguísticos marcados e não marcados, a complexidade formal e a frequência de uso estão numa dimensão mais concreta de aferição, perceptíveis no âmbito discursivo em que ocorrem. De outro modo, a complexidade cognitiva não é elemento de fácil constatação, pois requer uma definição mais criteriosa do que seria mais ou menos complexo do ponto de vista cognitivo. Como medir esforço mental e tempo de processamento quando se trata, por exemplo, das finais desenvolvidas e reduzidas? Numa análise rasa, dir-se-ia que as desenvolvidas são mais complexas cognitivamente por conterem verbo no modo subjuntivo e por terem, muitas vezes, a expressão de um sujeito não correferencial, o que lhe conferiria um *status* informacional de “novidade”. Mas não seriam as reduzidas relativamente complexas sob esse mesmo ponto de vista, já que mais inferências têm de ser feitas pelos interlocutores, em função de elas frequentemente terem sujeito não expresso, em grande parte por sua correferencialidade, além de estarem mais integradas sintaticamente à oração a que se vinculam? É um critério que parece exigir mais “critérios” para que se tome um caminho coerente de análise.

Tomando-se o princípio da marcação de forma mais ampla, não cabe estabelecer se um item é marcado ou não fora de qualquer contexto discursivo que possa ilustrar tal determinação, o que seria impertinente neste estudo, já que se pauta sobre dados da língua falada e escrita em uso. Entretanto, há exemplos na literatura linguística em que tal constatação se faz apenas pela intuição do linguista, como se pode ver no *Dicionário de linguagem e linguística* (TRASK, 2006, p. 187), com adaptações de exemplos à língua portuguesa:

O português *tigresa* é marcado com respeito a *tigre*, porque contém mais matéria morfológica e se aplica de maneira menos geral. O português *suspenso* é marcado com respeito a *suspendido*, porque tem uso limitado a certos contextos, e *mano* é marcado com respeito a *irmão*, porque carrega um significado emotivo extra, ausente na segunda forma. Uma sentença

passiva como *Janete foi detida pela polícia* é marcada com respeito à ativa correspondente *A polícia deteve Janete*, porque a passiva contém mais material, tem uma estrutura mais complexa e é mais rara do que a ativa.

Observa-se que parte dos exemplos é analisada com base na questão da complexidade formal e estrutural dos pares em comparação, como nos aspectos “mais material” e “estrutura mais complexa”, o que é inegável, porque é algo “palpável” e de fácil mensuração. Entretanto, apontar a raridade de uma forma em detrimento da outra (voz ativa/passiva) torna-se mais legítimo em termos de análise linguística quando são tomados os dados da língua em uso. Não apenas isso: o contexto discursivo tende a exercer um papel de grande relevância nesse sentido. Ou seja, a depender do gênero textual, seja da escrita ou da fala, ou ainda, dos graus de formalidade ou de informalidade em que a comunicação se efetiva, formas antes consideradas marcadas podem passar a ser não marcadas em outra ambiência discursiva, como pode ser o caso da distribuição formal das vozes do verbo anteriormente exemplificadas. Em Givón (1995, p. 27), encontra-se bastante presente a noção da importância contextual para o trabalho com a noção de marcação: “A marcação é um fenômeno que depende do contexto por excelência. Uma mesma estrutura pode ser marcada num contexto e não marcada em outro”.¹³

Numa abordagem funcionalista, a noção de marcação, além de depender dos contextos linguísticos em análise, também pode ser entendida e aplicada de forma escalar, gradiente. Desse modo, não necessariamente devem estar opostas apenas duas estruturas, sendo uma a marcada e a outra a não marcada. Pode haver entre elas níveis (graus): da mais para a menos marcada e vice-versa – tanto no que diz respeito aos critérios aqui já apontados, quanto aos contextos em que se encontram dispostas as formas linguísticas. Assim, na esteira do princípio da marcação considerado como escalar, surge a necessidade de uso do termo *protótipo*: “[...] as categorias são definidas pelo agrupamento de um número de características principais, as quais tendem a caracterizar o **protótipo**” (GIVÓN, 1995, p. 29).¹⁴ Ou seja, um modelo recorrente que se tem para determinado fenômeno linguístico com possibilidades variáveis. Por exemplo, ao final deste trabalho, na seção 5.5, serão propostas algumas escalas para as orações finais, com base no princípio da marcação, em termos da frequência de uso

¹³ Tradução livre a partir do seguinte fragmento do texto original: “Markedness is a context-dependent phenomenon par excellence. The very same structure may be marked in one context and unmarked in another”.

¹⁴ Tradução livre a partir deste fragmento do texto original: “[...] categories are defined by clustering of a number of central features, those that tend to characterize the **prototype**”

das variantes, e será visto que, associadas algumas características atribuídas ao sujeito da oração final, o protótipo (o modelo mais frequente de oração final) nos dados da fala e da escrita analisados é o mesmo.

Para ilustrar o que se vem dizendo, analisem-se alguns exemplos dos *corpora* que compõem o banco de dados deste estudo:

- (1) O mundo dá voltas. Precisou de séculos de opressão e de teorias machistas, de bruxas queimadas e sutiãs também, PARA QUE A MULHER RECUPERASSE O STATUS QUE JÁ TEVE UM DIA.
- (2) [...] eu tenho que gerenciar e organizar... fazer um esquema... PRA QUE TUDO FUNCIONE MUITO BEM [...]
- (3) Só que os físicos, cosmologistas e astrônomos não são pagos PARA TEREM BOM SENSO – sua obrigação é investigar o cosmo com todo rigor e descobrir do que ele é realmente feito, por mais estranho que possa parecer.
- (4) [...] então por isso que eu preciso que vocês me apoiem PARA COMPRARMOS mais brinquedos... mais jogos... mais balanços [...]

Os dados (1) e (3) são da língua escrita, e os dados (2) e (4), da língua falada. Em (1) e (2), em que a oração final aparece sob a forma desenvolvida (introduzida pela locução *para que* e com verbo expresso no modo subjuntivo), encontram-se, de acordo com os critérios propostos por Givón (1995), as finais marcadas em relação aos exemplos apresentados em (3) e em (4), porque têm maior complexidade formal – mais material – o que pode acarretar maior complexidade cognitiva também. Entretanto, se se opuserem os exemplos (3) e (4) – não marcados em relação a (1) e a (2), com verbo no infinitivo (flexionado) e apenas introduzidos pela preposição *para* – aos exemplos (5) e (6), abaixo relacionados, será verificado nos últimos um grau a menos no que diz respeito à marcação, pois estes apresentam a forma verbal ainda mais reduzida, sem flexão no infinitivo, ou seja, com uma complexidade formal menos saliente:

- (5) De um modo geral, as abordagens alternativas não existem PARA CURAR DOENÇAS, mas PARA PREVENI-LAS e PARA COMPLEMENTAR UM TRATAMENTO CONVENCIONAL.

(6) [...] meu pai já não podia... mais trabalhar... minha mãe... trabalhando direto... então eu tinha que:... estudar e:... logo arrumar alguma coisa PRA:: ... GANHAR DINHEIRO... eu ajudava meu pai ...

O que se quer mostrar é que pode haver escalaridade/gradiência no que diz respeito à marcação linguística, havendo níveis intermediários entre itens mais e menos marcados. Nesse sentido, sem ainda verificar o critério da frequência de uso, a categoria mais marcada é a oração final desenvolvida com sujeito explícito e não correferencial (exemplos 1 e 2); um pouco menos marcada é a final reduzida com sujeito não explícito e com verbo no infinitivo flexionado (exemplos 3 e 4); e a menos marcada é a final reduzida com sujeito não explícito correferencial e com verbo sem flexão no infinitivo (exemplos 5 e 6). Como exemplos de finais reduzidas com flexão de infinitivo encontrados foram raríssimos, será visto que essa propriedade não foi controlada para fins da análise variável do fenômeno finais desenvolvidas e reduzidas, servindo aqui como mera ilustração para uma definição de marcação linguística que se desenvolverá na seção 5.5 deste trabalho.

Como já mencionado, o princípio da marcação também deve estar atrelado aos contextos discursivos em que o fenômeno linguístico emerge. Em relação a dados da fala e da escrita – que compõem os *corpora* desta pesquisa –, Givón (1995, p. 30) afirma que a escrita formal é mais marcada em relação à oralidade informal porque apresenta, por exemplo, mais subordinação que coordenação – ou seja, maior complexidade estrutural –, além de apresentar aspectos morfológicos mais diversificados que a fala. A perspectiva adotada neste trabalho, no entanto, é a de verificar, em termos da frequência de uso, uma escala que vai da variante menos à mais marcada nos *corpora* falado e escrito, tendo em vista a convergência de alguns fatores estruturais. Desse modo, é preciso observar que os

[...] vários aspectos ligados ao caráter marcado, ou não marcado, dos enunciados envolvem grande complexidade e têm de ser examinados independentemente um do outro, de tal modo que a busca de correlação entre eles não invoque matéria de fé [...] mas repouse sobre verificação empírica (NEVES, 1997, p. 105).

Na próxima seção, são tratados outros dois princípios funcionalistas, o da economia e o da iconicidade, vistos na perspectiva da motivação linguística.

3.2.2 Economia e iconicidade

Tem sido recorrente na linguística contemporânea trabalhos na linha variacionista se utilizarem de princípios de orientação funcionalista na interpretação de determinados resultados encontrados. Dois desses princípios são a economia e a iconicidade – nos termos de Haiman (1983), entendidos como motivações. Particularmente no Brasil, e numa abordagem teórica mais ampla, trabalhos como os de Braga (1996), Scherre (1997) e Furtado da Cunha (2001), apenas para citar alguns, valeram-se desses dois princípios, entendidos muitas vezes como forças em competição (DU BOIS, 1985), na busca pela compreensão dos fenômenos linguísticos de suas pesquisas, mas chegaram a conclusões divergentes: a não aplicação desses princípios ao fenômeno em investigação – Scherre (1997); a confluência dessas motivações na compreensão do fenômeno, sem que estivessem em competição – Braga (1996); e a constatação da competição entre as duas forças – Furtado da Cunha (2001).

Nesse contexto, deve-se entender melhor o que é economia e iconicidade em termos linguísticos. De forma generalizada, “a economia é o princípio segundo o qual as expressões devem ser minimizadas sempre que possível” (CROFT, 1990, p. 102)¹⁵. Haiman (1983, p.802) diz que “[...] quanto mais imprevisível for uma informação, mais codificação ela requer”¹⁶. O inverso dessa relação deve ser assim depreendido: mais previsibilidade informacional implicaria menos codificação linguística.

Assim, aplicando a ideia de motivação econômica ao fenômeno da presente análise, pode-se atribuí-la a alguns contextos, como quando nas variantes se observa o sujeito expresso ou não expresso em sua estrutura sintática. Nos dados da escrita,

(1) Na natureza, a mãe corta essa ligação PARA QUE O FILHOTE APRENDA A CUIDAR DE
SUA VIDA [...]

e

¹⁵ Tradução livre do original: “Economy is the principle that the expressions should be minimized where possible” (CROFT, 1990, p. 102).

¹⁶ Tradução livre do original: “[...] the more unpredictable a piece of information is, the more coding requires” (HAIMAN, 1983, 802)

(2) Mas por que cada vez mais pessoas abandonam suas religiões PARA TORNAREM-SE EVANGÉLICAS?

os sujeitos das orações finais se comportam de forma distinta. Em (1), o sujeito “o filhote” não tem correferência na oração anterior, cujo sujeito é “a mãe”, o que motiva a explicitação do primeiro na final desenvolvida, para que o propósito comunicativo seja efetivado. Já em (2), o sujeito da final reduzida, “as pessoas”, é correferente ao da oração principal e não se explicita. Entende-se, portanto, que a não correferência em (1) e a correferência em (2) tendem a motivar respectivamente a explicitação e a não explicitação do sujeito nas orações finais, uma vez que na primeira a informação é imprevisível e requerente de codificação, e na segunda é previsível, não necessitando de codificação (realização). Nesse sentido, sujeito expresso não correferente tende a fazer emergir discursivamente finais desenvolvidas, enquanto sujeito não expresso correferente é um elemento que tende a favorecer a ocorrência de orações finais reduzidas no discurso – essas e outras combinações nesse sentido serão mais bem discutidas ao longo do capítulo 5 deste trabalho.

Pelos exemplos anteriores, verificou-se que a motivação econômica age na estrutura interna das orações finais no que diz respeito ao seu sujeito. Assim, já que o desenvolvimento e a redução das orações finais estão sendo tomados como dependentes da influência de outros fatores, linguísticos ou não, cabe investigar se há outra força motivadora aí envolvida que não a economia. O raciocínio é o seguinte: por que mais codificação tende a promover o uso da final desenvolvida, e menos codificação o da final reduzida? Para uma tentativa de resposta a essa pergunta, veja-se um pouco a respeito do princípio da iconicidade nos domínios linguísticos.

Um pouco distante do isomorfismo de Bolinger (1975), para quem há uma relação mais estreita de motivação entre forma e função nas línguas – uma forma para cada função –, o princípio da iconicidade se enquadra nos estudos funcionalistas atuais de maneira menos determinística.

A iconicidade tem a ver com uma relação motivada entre o pensamento e seu reflexo na realidade externa (GIVÓN, 1995). É um princípio do funcionalismo que se opõe à relação de arbitrariedade na linguagem humana, como se apregoava no chamado polo formalista dos estudos linguísticos. Em Saussure (1975), encontra-se uma das principais referências à arbitrariedade das línguas humanas nos estudos linguísticos modernos. Ao tratar da relação

entre significante e significado, o autor diz que “[...] o significante é *imotivado*, isto é, arbitrário em relação ao significado, com o qual não tem nenhum laço natural na realidade” (SAUSSURE, 1975, p.83).

Nos estudos contemporâneos, no entanto, essa visão é cada vez mais refutada, e uma relação motivada da realidade externa com as manifestações linguísticas ganha cada vez mais apoio entre os pesquisadores, o que contribui para a flexibilização tanto da relação arbitrária quanto da isomórfica. É uma visão propiciada por “estudos sobre os processos de variação e mudança, ao constatar a existência de duas ou mais formas alternativas de dizer ‘a mesma coisa’” (FURTADO DA CUNHA, 2001, p. 4-5).

Assim, do ponto de vista funcional, a iconicidade pode ser entendida como uma relação motivada entre estruturas morfossintáticas e suas funções semânticas ou pragmáticas correspondentes. A noção de motivação considera o fato de que o usuário da língua, de alguma maneira, busca estabelecer uma relação da forma (da palavra, no âmbito da morfologia; da sentença, no âmbito da sintaxe) com o significado expresso.

O princípio da iconicidade tem sido aplicado aos estudos linguísticos a partir dos seus subprincípios: (a) *o subprincípio da quantidade*; (b) *o subprincípio da ordenação sequencial*; e (c) *o subprincípio da integração ou da proximidade*.

Na busca por tentar responder à pergunta feita na página anterior – por que mais codificação tende a promover o uso da final desenvolvida, e menos codificação, da final reduzida? –, recorre-se ao primeiro subprincípio, o da quantidade. Assim, como a oração final com sujeito expreso não correferencial é quantitativamente mais extensa, a relação icônica reflete uma codificação oracional também maior, nesse caso a variante final desenvolvida. De igual modo, a oração final com sujeito não expreso correferencial tem uma menor extensão que se reflete no uso mais recorrente das finais reduzidas, daí a relação icônica, ou seja, se há menos codificação, há mais redução.

É importante lembrar que, ao selecionar a variante final desenvolvida, a maior codificação que o usuário terá de utilizar consiste no emprego de locução conjuntiva, em geral *para que*, de sujeito explícito, na maioria das vezes, e de verbo no modo subjuntivo, como se pode constatar no exemplo (1), aqui retomado como (3):

- (3) Na natureza, a mãe corta essa ligação PARA QUE O FILHOTE APRENDA A CUIDAR DE SUA VIDA [...]

Processo sintático “inverso” ocorre no uso da variante final reduzida, cuja estrutura prototípica, em termos empíricos, costuma ter a preposição *para*, sujeito não explícito, por ser correferencial (atuação da motivação econômica), e verbo no infinitivo.

Em se confirmando a influência dos fatores ligados ao sujeito, (não) correferencialidade e (não) explicitação, na variação sintática das orações adverbiais finais – confirmação esta que será buscada por meio dos recursos estatísticos dos quais se falará no capítulo 4 –, a expectativa é de que nela estejam envolvidas duas forças motivadoras: uma, a econômica, manifesta em sua estrutura interna, na função sintática de sujeito; a outra, a icônica, no reflexo quantitativo da codificação no uso variável da final desenvolvida ou reduzida.

Além do subprincípio icônico da quantidade, que é o que mais interessa a este trabalho, há também o da ordenação sequencial, de acordo com o qual as informações no nível discursivo são dispostas da seguinte maneira: as que constituem o tópico vêm, em geral, primeiro no discurso, e as de comentário vêm depois, numa relação de ordem entre informações “velhas” e “novas”. Assim, na relação entre a oração principal e a adverbial final, a tendência da ordenação sequencial ou linear é a da posposição da oração final em relação à principal, predominância que se confirmará ou não a partir da análise dos dados da língua em uso. Isso pode ser entendido como um reflexo da ordem como os fatos ocorrem na realidade externa na expressão da finalidade.

O terceiro subprincípio da iconicidade, o da integração ou proximidade, revela que “[...] o conteúdo que está mais próximo cognitivamente também estará mais próximo sintaticamente, ou seja, conceitos que estão mentalmente mais integrados tendem a ser codificados mais próximos na corrente linguística” (MARTELOTTA, 2006). Assim, esse princípio se aplicaria, por exemplo, aos casos de redução das finais, como no seguinte dado da língua falada:

- (4) [...] ela vai lá PRA SABER COMO É QUE TÁ e volta pra falar dela.

Neste dado, um grau mais alto de integração entre o que se veicula na oração principal e na final é revelado pela estrutura reduzida e pela identidade dos sujeitos também. Aqui, na verdade, pode-se entrever a atuação também dos dois primeiros subprincípios arrolados: o da

quantidade (menos informação, menor estrutura linguística); e o da ordenação sequencial (final posposta).

Trabalhos na linha dos graus de integração entre orações (GÖRSKI, 2000, 2001) apontam um forte laço nesse sentido entre as finais e as orações às quais estão vinculadas, o que significaria que os fatores anteriormente mencionados, como a ausência de sujeito explícito e sua correferencialidade, contribuiriam para aumentar o grau de integração entre a oração principal e as finais reduzidas, que, na posição posposta, estariam num estágio mais gramaticalizado na língua. Contudo, é importante ressaltar que averiguar níveis de integração entre as orações não constitui tarefa desta dissertação.

Em suma, entre os subprincípios da iconicidade, o que mais interessa a este trabalho é o da quantidade, por estar relacionado à primeira motivação, a econômica. Nesse sentido, retomando o que se disse no início desta seção, a constatação que se faz é de que as duas motivações não estão em competição no fenômeno de variação sintática das adverbiais finais, mas ambas são úteis, assim como em Braga (1996), para o entendimento de parte desse fenômeno.

3.3 Revisão bibliográfica

Para que se desenvolvesse de forma mais aprofundada esta pesquisa, alguns trabalhos de natureza acadêmica, como teses, dissertações e artigos, foram consultados. Nesse sentido, esta breve seção vai apresentar alguns aspectos desses trabalhos, em virtude de o objeto de estudo investigado e a linha teórico-metodológica adotada nesta dissertação serem semelhantes aos encontrados nesses estudos.

Azevedo (2000), em tese de doutorado, buscou identificar as formas de manifestação da expressão da finalidade em português. Para isso, trabalhou especificamente com a língua escrita, na modalidade padrão. O autor utilizou como referência teórica o funcionalismo norte-americano, trabalhando com princípios como gramaticalização e marcação. Lançou mão também da quantificação de dados, utilizando para tanto o programa computacional Varbrul. Quanto ao objeto de estudo, não se restringiu apenas às adverbiais finais, trabalhando com outras formas de expressão da finalidade, oracionais ou não. Seu estudo pôde demonstrar que os princípios teóricos do funcionalismo foram instrumentos de análise importantes para a

compreensão de que o usuário da língua não “escolhe” apenas estilisticamente como vai expressar a finalidade na língua portuguesa, mas que tal expressão está ligada a fatores cognitivos e a necessidades semânticas e pragmáticas também. No capítulo 5, alguns dos resultados encontrados por Azevedo serão comparados, oportunamente, aos desta dissertação.

Buscando compreender os mecanismos de expansão e redução oracionais, Finck (2000) utilizou parte do *corpus* do projeto VARSUL – Variação Linguística Urbana na Região Sul –, correspondente à cidade de Florianópolis. Um estudo focado na língua falada, cujo recorte se fez entre as orações adverbiais finais, temporais e causais. Alinha-se teoricamente ao funcionalismo norte-americano e utiliza a Teoria da Variação para tratar de prováveis fatores linguísticos ou não que possam atuar nos mecanismos de expansão e redução dessas orações. A autora trabalhou na perspectiva de verificar graus de integração no nível oracional, bem como perceber estágios de gramaticalização de determinados tipos de orações. Os resultados obtidos foram, de forma generalizada, os seguintes: as finais tendem a ocorrer predominantemente como reduzidas; as causais se apresentaram mais frequentemente como desenvolvidas, e as temporais foram as que se comportaram de forma mais variável entre uma e outra variante. Também no capítulo 5, alguns dos resultados encontrados no trabalho de Finck serão comparados aos obtidos nesta dissertação.

Também numa perspectiva funcionalista, na linha norte-americana, Dias (2001) desenvolveu um trabalho de pesquisa sobre as cláusulas de finalidade em português, não se restringindo às adverbiais finais. Trabalhou com a língua falada, com base no banco de dados do PEUL, do Rio de Janeiro, e também com textos escritos de natureza jornalística. A autora buscou tratar das formas marcadas e não marcadas em cada ambiência discursiva e chegou a algumas constatações: a forma típica não marcada é a reduzida, cuja estrutura é representada por *para* + *infinitivo*, ocorrendo na posição posposta à oração matriz. Faz ainda uma série de subcategorizações, como as cláusulas de finalidade parentéticas e de adendo, que estão relacionadas às condições dos atos de fala. É um estudo bastante fundamentado em questões discursivas e situacionais.

A mesma autora, (DIAS, 2002), publicou artigo no qual relaciona as convergências e divergências das chamadas cláusulas de finalidade na fala e na escrita, sobretudo no que concerne à posição que elas ocupam em relação às orações núcleos. Segundo a autora, dadas as finais antepostas, mediais e pospostas, as primeiras “[...] ocorrem com mais frequência na escrita para auxiliar o leitor no desenvolvimento do fluxo informacional” (DIAS, 2002, p.

532). Na mesma publicação, a pesquisadora também descreveu de forma sucinta o preenchimento da posição de sujeito nas orações às quais as finais estão vinculadas.

Em tese de doutorado, Menezes (2001) definiu seu objeto de pesquisa como as “construções” infinitivas iniciadas por *para*. Para tanto, baseou-se na língua falada e escrita e trabalhou com a noção de graus de redução das orações infinitivas na língua portuguesa. Procurou identificar, a partir da noção de protótipo, as construções mais e menos reduzidas, por meio de um tratamento escalar do processo de redução. A autora propõe três graus de redução, utilizando a designação *construções abertas e fechadas*, com alguns níveis intermediários. Assim, as orações infinitivas definidas como no grau 1 são as menos reduzidas, estando mais próximas às orações desenvolvidas. As em grau 2 estão no ponto intermediário, enquanto as em grau 3 seriam as mais reduzidas, estando numa linha fronteira entre oração e não oração (com características mais nominais).

Thompson (1985), em trabalho com as orações finais no inglês escrito, pondera a respeito da posição desse tipo de oração, dizendo que a anteposição e a posposição delas são iconicamente motivadas. Ou seja, a posposição motivada por iconicidade está vinculada à ideia de que se declara uma “ação” para depois se dizer do objetivo/finalidade que se tem a partir dela. Enquanto a motivação icônica da anteposição reside no fato de que, em se tratando de discurso escrito, as finais antepostas orientam discursivamente a atenção do leitor, uma vez que as orações que iniciam o período, quando se referem a um problema já dito no discurso, teriam de se colocar antes da oração principal, na qual se encontraria uma possível saída para esse problema.

No capítulo seguinte, passa-se a apresentar os aspectos metodológicos que guiaram o desenvolvimento desta pesquisa, tais como os *corpora* selecionados e o levantamento dos dados, uma breve descrição do programa computacional Varbrul e a definição das variáveis independentes utilizadas na análise das ocorrências das orações adverbiais finais desenvolvidas e reduzidas.

CAPÍTULO 4

METODOLOGIA

4.1 Os *corpora* em análise e a seleção dos dados

Esta pesquisa na área dos estudos analítico-descritivos da linguagem, além de buscar cercar-se de um aparato teórico que embase as suas reflexões linguísticas (conforme visto no capítulo anterior), selecionou dois *corpora* para que se fizesse a coleta dos dados que pudessem ilustrar o fenômeno de variação sintática entre as finais desenvolvidas e reduzidas. Um dos *corpora* é da língua falada, e o outro é da língua escrita do português brasileiro.

O *corpus* da língua falada é composto de um conjunto de entrevistas da base de dados do PORTVIX – designação do projeto *O português falado na cidade de Vitória*, desenvolvido pelo NUPLES / Ufes – Núcleo de Estudos Linguísticos do Espírito Santo – com o objetivo de “[...] explicitar os hábitos linguísticos, da comunidade de fala em questão, revelando suas particularidades e suas confluências com outras comunidades, especialmente as mais próximas, no caso, a do Rio de Janeiro e de Minas Gerais” (YACOVENCO, 2002, p. 102-103).

A justificativa para a escolha desse *corpus* se faz, em primeiro lugar, por ser um legítimo exemplo de entrevistas nos moldes labovianos¹⁷, que leva em consideração fatores sociais na seleção dos entrevistados, e também pela própria técnica de entrevistar, direcionando os assuntos e temas para que o informante se sinta à vontade e possa falar de maneira descontraída e menos controlada, a fim de que a língua possa ser observada em seu funcionamento e uso não monitorados.

Assim, em números reais, foram coletados dados de 19 (dezenove) das 46 (quarenta e seis) entrevistas que compõem esse banco de dados, perfazendo um total de 302 (trezentas e duas) ocorrências de orações subordinadas adverbiais finais desenvolvidas e reduzidas. O fato de não ter sido analisada toda a amostra PORTVIX se justifica no sentido de que as ocorrências

¹⁷ O que se quer dizer é que essa ideia está diretamente vinculada ao referencial teórico da sociolinguística variacionista, cuja base se assenta principalmente nos estudos de Labov (2008), como se viu no capítulo 3.

estavam se padronizando à medida que se avançava na coleta dos dados; isto é, era encontrada uma massa grande de dados de finais reduzidas e pequena de desenvolvidas. Além disso, numa busca aleatória em mais 8 entrevistas, não foi encontrado nenhum dado de final desenvolvida, o que reforçou a decisão de se trabalhar apenas com as 19 primeiras células da base de dados referida, que já haviam sido rastreadas.

Quanto ao *corpus* da língua escrita, este é formado por um conjunto de reportagens veiculadas pela revista *Superinteressante*¹⁸, cuja característica central é o jornalismo de natureza científica. Essas reportagens foram veiculadas entre o mês de junho de 2003 e junho de 2004. Tal recorte temporal foi implementado para que o *corpus* da escrita fosse contemporâneo ao das entrevistas do PORTVIX.

Escolher analisar o fenômeno de variação sintática das orações adverbiais finais também na língua escrita se justifica pelo fato de se esperar encontrar diferenças no emprego das desenvolvidas e reduzidas em relação à fala, porque a escrita, em especial a do gênero textual que se elegeu, tem um caráter mais formal que a fala das entrevistas, o que cria a expectativa de que pelo menos a frequência nos usos possa apresentar alguma nuance interessante para a compreensão e a interpretação do fenômeno em questão.

Foram extraídos dados de 35 (trinta e cinco) reportagens, o que resultou num total de 282 (duzentas e oitenta e duas) ocorrências de orações subordinadas adverbiais finais, entre desenvolvidas e reduzidas.

Foram selecionados, preliminarmente, tanto das entrevistas quanto das reportagens, todos os casos que contivessem as características típicas das adverbiais finais: (1) para (pra) que + verbo no modo subjuntivo; (2) para (pra) + verbo no infinitivo (flexionado ou não). Em seguida foi selecionado apenas o tipo de estrutura que constitui o presente objeto de estudo. As outras construções que estão fora do campo de observação desta pesquisa, mas que conservam traços estruturais semelhantes aos das adverbiais finais, e encontradas nos *corpora* aqui descritos são, entre outras,

(a) as substantivas objetivas diretas:

¹⁸ Periódico publicado pela Editora Abril, pertencente ao Grupo Abril, empresa brasileira com sede na cidade de São Paulo – SP. Além da *Superinteressante*, a editora tem outras publicações de grande circulação nacional, como a revista *Veja*.

“toda vez que tem... algum evento lá que eles fazem que eles fazem eventos lá aí a gente sempre... eles pede sempre A GENTE PRA TOCAR LÁ a gente toca lá ... sem cobrar nada em troca”;

(b) as substantivas completivas nominais:

“bom professor é aquele que que sabe explicar que acompanha o raciocínio do aluno... sem que ele tenha dificuldade PRA APRENDER A MATÉRIA e por aí assim vai”;

(c) as adverbiais de frase ou de enunciado:

“eu penso que... não são... relativamente eles são... eu PRA DIZER A VERDADE... não ando mais... não ando mais de ônibus...”

(d) as construções dar + para + infinitivo:

“tem uma bateria no meu quarto ali... toco aí de vez em quando não dá PRA TOCAR a semana toda”¹⁹

Em muitos casos, quando persistia a dúvida entre uma ou outra ocorrência, a técnica utilizada para a confirmação de se tratar de uma adverbial final e não de uma substantiva completiva nominal foi a da clivagem e da focalização (NEVES, 2000, p. 890) – conforme descrito no capítulo 2. Isso se demonstra em (e) que, clivada e focalizada em (f), tem atestada a sua natureza circunstancial:

(e) “[...] ele passa por uma seleção PARA SABER EM QUE CASA INGRESSARÁ: é submetido ao Chapéu Seletor, que avalia os novatos e os destina, por afinidade, a uma delas.”²⁰

(f) [...] *É PARA SABER EM QUE CASA INGRESSARÁ* *que* ele passa por uma seleção: é submetido ao Chapéu Seletor, que avalia os novatos e os destina, por afinidade, a uma delas.

¹⁹ Os exemplos de (a) a (d) são do PORTVIX.

²⁰ Exemplo do *corpus* da língua escrita.

Na seção seguinte, será descrito brevemente o procedimento para o tratamento quantitativo dos dados.

4.2 O Varbrul e as variáveis independentes

A opção por trabalhar com um fenômeno linguístico de natureza variável requer um instrumental de análise que possa dar conta da proporção com que cada fator, interno ou externo ao sistema linguístico, contribui para o uso mais frequente de uma ou de outra variante da variável dependente. Assim, “O problema central que se coloca para a Teoria da Variação é a avaliação do *quantum* com que cada categoria postulada contribui para a realização de uma ou de outra variante das formas em competição” (NARO, 2008, p. 16). Naro ainda acrescenta que a atuação do conjunto de fatores que influencia a alternância entre as variantes em competição se dá de forma concomitante. Então, o desafio é “[...] isolar e medir separadamente o efeito de um fator [...] quando tal fator nunca se apresenta isoladamente nos dados” (Id., p. 17).

Para que houvesse condições mais precisas que dessem conta desse *quantum*, optou-se pelo uso do programa computacional de tratamento quantitativo-estatístico Varbrul, na sua versão para o ambiente Windows, denominada GOLDVARB X (SANKOFF; SMITH; TAGLIAMONTE, 2005), que

[...] é um conjunto de programas computacionais de análise multivariada, especificamente estruturado para acomodar dados de variação sociolinguística. A análise se chama ‘multivariada’ porque permite investigar situações em que a variável linguística em estudo é influenciada por vários elementos do contexto, ou seja, múltiplas variáveis independentes. A investigação mede os efeitos, bem como a significância dos efeitos dessas variáveis independentes sobre a ocorrência das realizações da variável que está sendo tratada como dependente. (GUY; ZILLES, 2007, p. 105)

O que Guy e Zilles (2007) chamam de *variável dependente* é o que neste trabalho se selecionou como fenômeno para a análise: a variação sintática das orações subordinadas adverbiais finais, desenvolvidas ou reduzidas. E as *variáveis independentes* são os fatores pré-selecionados pelo pesquisador como aqueles que podem ter influência na variação do fenômeno linguístico em estudo.

Nesse sentido, as variáveis independentes (também chamadas de *grupos de fatores*) definidas nesta pesquisa para melhor entender a alternância entre finais desenvolvidas e reduzidas estão divididas entre as linguísticas e as sociais. Desse modo, as variáveis independentes linguísticas são:

(1) *Correferencialidade do sujeito da adverbial final*

Esta variável independente está subdividida em três fatores, a saber: (a) sujeito correferente ou igual; (b) sujeito parcialmente correferente; (c) sujeito não correferente ou diferente. Essa relação é verificada entre os sujeitos das orações principal e subordinada final. No caso do sujeito parcialmente correferente, o sujeito codificado na oração final é a retomada, explícita ou não, de algum elemento presente na oração à qual está vinculada, desde que não seja o sujeito desta. A escolha desta variável está diretamente ligada ao fato de nos primeiros dados levantados a correferencialidade entre os sujeitos parecer indicar maior índice de redução que de desenvolvimento da oração final, do mesmo modo que a percepção contrária também era verificada, ou seja, a ausência da correferência nessa função sintática propiciava mais casos de finais desenvolvidas. Além disso, a correferencialidade está intimamente ligada à explicitação do sujeito na oração final, aspecto que está relacionado no item 4, abaixo. Alguns trabalhos anteriores a este, como o de Finck (2000), a partir do qual se decidiu controlar a correferencialidade parcial do sujeito, e Görski (2001), avaliaram a relação existente entre correferencialidade e explicitação do sujeito nas orações finais. Neste trabalho, a expectativa é poder confirmar as primeiras tendências do levantamento dos dados – a partir também das conclusões das pesquisas mencionadas –, verificando se há influência, do ponto de vista estatístico, dos fatores que compõem esta variável e em que medida eles favorecem uma ou outra variante do fenômeno em estudo, para que, posteriormente, a análise seja feita do ponto de vista linguístico.

São exemplos do PORTVIX para os três fatores aqui definidos:

(a) *sujeito correferente*: [...] meu tio que trabalha na... vai vir aqui pra... PRA CONSERTAR O COMPUTADOR [...]

Neste exemplo de final reduzida, o sujeito da oração final é *meu tio*, que está expresso na oração principal, daí a sua correferencialidade.

(b) *sujeito parcialmente correferente*: [...] em vez disso eles poderia soltar os peixe e eles não soltam aí acaba matando aí em vez de soltar os peixes PRA CRESCER E PODER REPRODUZIR MAIS [...]

Neste caso de final reduzida, o sujeito da oração final é *os peixes*, que na oração anterior funciona sintaticamente como objeto direto do verbo *soltar*, e é diferente do sujeito *eles*. O rótulo *sujeito parcialmente correferente* foi uma tentativa de diferenciar esse tipo de ocorrência daqueles em que o sujeito é totalmente não correferente, não retomando nem o sujeito e nem qualquer outro termo da oração à qual a final esteja vinculada, como se vê no próximo exemplo.

(c) *sujeito não correferente*: [...] eu tenho que gerenciar e organizar... fazer um esquema... PRA QUE TUDO FUNCIONE MUITO BEM.

Este dado de final desenvolvida traz o sujeito *tudo* como não correferente a nenhum termo da oração anterior, nem ao sujeito, nem ao objeto direto.

A seguir, vêm as especificações da segunda variável independente.

(2) Características sintático-semânticas do sujeito da oração principal

Por se entender que a oração à qual se liga a adverbial final apresenta prototipicamente um sujeito que controla o que se passará na expressão de finalidade (NEVES, 2000, p. 888), neste grupo foram especificados dois fatores para que se verifique se, quando o tipo de sujeito não é o esperado pela forma mais típica das finais, há algum tipo de influência na variação entre finais desenvolvidas e reduzidas. As primeiras expectativas são de que as estruturas com sujeito não controlador ou oracional, quando presentes na oração vinculada à final, parecem propiciar a emersão discursiva das finais desenvolvidas – em virtude das primeiras observações dos dados, sobretudo na língua escrita. Quanto ao sujeito prototípico, devido ao equilíbrio encontrado nas primeiras ocorrências analisadas, não há uma expectativa muito clara e definida no que tange ao favorecimento a uma ou a outra variante em estudo. É o que vai ser averiguado, portanto. Desse modo, os fatores deste grupo são: (a) sujeito controlador/animado; (b) sujeito não controlador/inanimado/oracional, ilustrados a seguir com exemplos da escrita.

(a) *sujeito controlador/animado*: Desde que o mundo é mundo, as pessoas vão aos "médicos" [curandeiros, pajés, xamãs...] PARA FALAR DE SUAS ANSIEDADES, SUAS DORES E OBTER ALGUM CONFORTO - nem sempre a cura.

Verifica-se neste exemplo de final reduzida que o sujeito da oração principal, *as pessoas*, desenvolve a ação de ir aos “médicos” de forma a delimitar, “controlar” o que se passará posteriormente, que é falar de suas ansiedades, dores, etc.

(b) *sujeito não controlador/inanimado/oracional*: Um exemplo cotidiano da facilidade feminina em criar intimidade são as relações que mulheres desenvolvem no salão de beleza. Bastam 15 minutos PARA QUE ELAS COMECEM A FAZER CONFISSÕES AO SEU CABELEIREIRO.

Neste dado de final desenvolvida, ocorre na oração principal um tipo de sujeito não controlador e ao mesmo tempo inanimado, já que *15 minutos* constitui o sujeito de *bastam*. Essa estrutura tem comportamento semelhante ao do sujeito oracional, como se vê no seguinte exemplo, também com o verbo *bastar*:

[...] Depois porque, como acontece com o gelo, basta reaquecer o vácuo PARA QUE A MATÉRIA-ESQUELETO SAIA DO SEU ABISMO PARA O ANDAR DE CIMA DO UNIVERSO.

Passa-se, agora, à terceira variável independente e a seus fatores.

(3) *Voz/aspecto semântico do verbo da adverbial final*

A justificativa para a escolha deste grupo de fatores para a análise está no fato de que, apenas seguindo a intuição linguística, parece ser que as construções na voz passiva ou com verbos que indicam passividade do sujeito da oração final tendem a favorecer a ocorrência da final desenvolvida e desfavorecer a final reduzida. Assim, para averiguar se a voz do verbo ou mesmo seu aspecto semântico podem ter alguma influência sobre a alternância entre a final desenvolvida e a reduzida, fez-se a subdivisão desta variável nos seguintes fatores: (a) voz ativa / verbo que expresse o sujeito como agente, (b) voz passiva / verbo que expresse o sujeito como paciente. Eis os exemplos para cada fator, extraídos da base de dados da fala:

(a) *voz ativa / verbo que expresse o sujeito como agente*: [...] ele bota muita força no braço PRA PODER MOVIMENTAR porque ficar trazendo aquilo tudo é um peso, entendeu [...]

Percebe-se neste exemplo de final reduzida que a locução verbal *poder movimentar* indica o sujeito, *ele*, como agente do que se declara no discurso, enquanto no exemplo a seguir, de

final desenvolvida, há um sujeito paciente, *os privilégios do governo*, atestado pela estrutura de voz passiva, com o tempo composto *sejam dados*.

(b) *voz passiva / verbo que expresse o sujeito como paciente*: [...] então acho que tem que ser... tem que ter uma forma melhor de/ de analisar de avaliar PRA QUE REALMENTE [...] OS PRIVILÉGIOS DO GOVERNO SEJAM DADOS A QUEM NECESSITA.

(4) *Explicitação do sujeito na adverbial final*

Este fator refere-se à explicitação ou não do sujeito na subordinada. É um item de grande relevância, pois traz à cena analítica a questão da funcionalidade comunicativa da explicitação ou do cancelamento do sujeito na adverbial final. Em primeiro lugar, espera-se verificar se, de forma isolada, a expressão ou não do sujeito exerce algum papel de relevância estatística na variação entre as finais. Em se confirmando a tendência percebida nos primeiros dados coletados e em outros trabalhos já realizados (cf. item 1), a expectativa é de que a explicitação aponte favorecimento à ocorrência das finais desenvolvidas, e que a não explicitação à das reduzidas. Ademais, como se mencionou no item (1), sobre a correferencialidade, outro objetivo é buscar, numa análise posterior, a potencial imbricação sintática que há entre essas duas variáveis. São os fatores deste grupo: (a) sujeito explícito; (b) sujeito não explícito ou sujeito zero. Exemplos da língua escrita:

(a) *sujeito explícito*: Portanto, da próxima vez que uma coincidência acontecer, não se espante. Você provavelmente trabalhou PARA QUE ELA OCORRESSE.

(b) *sujeito não explícito*: O governo cobrava impostos altos sobre produtos manufaturados PARA PROTEGER A NASCENTE INDÚSTRIA DO NORTE DO PAÍS.

Em (a), verifica-se a explicitação do sujeito – *ela* – na final desenvolvida, enquanto em (b), final reduzida, sua expressão não ocorre. É digna de observação a notável imbricação que há entre esta variável e a correferencialidade do sujeito da oração final – conforme já apontado no capítulo anterior, em termos funcionais. Em (a), o sujeito exposto não tem correferência na oração anterior (há, é claro, correferência no discurso), e a oração se desenvolveu; em (b), o sujeito não exposto é correferencial, e a final se reduziu.

A seguir, a última variável independente de caráter linguístico.

(5) *Posição da adverbial final*

Pelo fato de a mobilidade da posição – anteposta ou posposta – ser considerada como um traço próprio das orações adverbiais, neste caso das finais, este grupo de fatores foi constituído para a análise no fenômeno de variação sintática que aqui se investiga. Chafe (1988, p.19) assevera que essa é uma das principais características que diferenciam as orações adverbiais das coordenadas: “Aparentemente, é esta a capacidade que se pode chamar de conexão bidirecional que distingue essas orações subordinadas das (orações coordenadas)”.²¹ Além disso, de acordo com Braga (2005, p. 264), as orações coordenadas, “[...] por codificarem iconicamente a realidade a que remetem, não admitiriam alterações na ordem”. Essa característica, no entanto, não se aplica às adverbiais, que, resguardados os casos em que ocorrem alterações semânticas, podem apresentar-se antepostas ou pospostas à oração núcleo. É preciso dizer que é recorrente em trabalhos que lidam com o estudo e análise de orações, seja na esfera da coordenação ou da subordinação, controlar a variável posição; recorrência esta que reforça o fato de este grupo de fatores ter sido escolhido para compor o quadro analítico desta pesquisa.

Assim sendo, os fatores deste grupo são: (a) final posposta; (b) final anteposta, controlados na expectativa de averiguar se a posição em que se encontra a oração final pode exercer alguma influência significativa na ocorrência de desenvolvidas e reduzidas. Num primeiro momento, somente se pôde constatar uma predominância de posposição, tanto na fala quanto na escrita, sem que se percebesse, de forma empírica, algum tipo de favorecimento à desenvolvida ou à reduzida. Seguem os exemplos correspondentes, sendo o primeiro da fala, e o segundo, da escrita:

(a) *final posposta*: [...] eles tem que estudar... aí amanhecer o dia lá, assim... tomando conta de carro do outros PRA GANHAR MIXARIA NÃO DÁ, meu pai não deixa não, meu pai não deixa eles trabalhar não [...]

(b) *final anteposta*: PARA FINANCIAR PROGRAMAS SOCIAIS COMO ESSES, os governos tiveram que aumentar a arrecadação.

Terminada a especificação das variáveis independentes linguísticas, devem ser conhecidas as restantes, de natureza social.

²¹ Tradução livre a partir do original: "Apparently it is this capacity for what might be called bidirectional linking that distinguishes these 'subordinate' clauses from [...] (coordinate clauses)."

No que tange aos aspectos sociais deste estudo, foram selecionadas 3 (três) variáveis independentes, em função dos critérios do *corpus* PORTVIX. Esses três grupos, bem como os fatores que os integram, são os seguintes:

(6) *Sexo/gênero*: (a) homem; (b) mulher.

(7) *Faixa etária*: (a) 7 a 14 anos; (b) 15 a 25 anos; (c) 26 a 49 anos; (d) 50 anos ou mais.

(8) *Nível de escolarização*: (a) fundamental; (b) médio; (c) superior.

Como se viu na seção 1.2 desta dissertação, não há expectativas significativas a respeito das possíveis influências que as variáveis sociais possam ter sobre o fenômeno de variação sintática das orações finais. Excetuando-se dessa consideração o nível de escolarização dos usuários, sobre o qual se hipotetizou poder haver alguma relação entre mais escolaridade e mais emprego de finais desenvolvidas.

Com base nessas variáveis independentes de caráter social, elaborou-se no quadro abaixo a discriminação dos informantes do PORTVIX, de acordo com as células das entrevistas que foram utilizadas:

QUADRO 1
Distribuição social dos informantes do *corpus* da fala

Faixa etária	7-14		15-25		26-49		50 OU +		
Sexo/gênero	H	M	H	M	H	M	H	M	
Ensino fundamental	1	1	1	1	-	1	1	1	=7
Ensino Médio	-	-	1	1	1	1	1	1	=6
Ensino superior	-	-	1	1	1	1	1	1	=6
Total de informantes									=19

Conforme já se antecipou na seção anterior o motivo da utilização de 19 entrevistas das 46 existentes no PORTVIX, este quadro serve para mostrar que se tentou buscar certo equilíbrio na coleta dos dados, em confluência com os três fatores sociais em questão. Pode-se avaliar que apenas 1 informante por célula não seja suficiente para uma análise confiável; entretanto, como também já se frisou, dada a consideração de que o objeto de estudo deste trabalho tem em seu viés linguístico maior relevância que propriamente no social, acredita-se que essa

distribuição não tenha afetado de forma negativa a pesquisa, cujos resultados serão apresentados no próximo capítulo.

No que diz respeito à base de dados da escrita, a distribuição se deu da seguinte maneira:

QUADRO 2
Quantificação dos jornalistas distribuída nos grupos de fatores sociais

Sexo/gênero	Faixa etária				
	7-14	15-25	26-49	50 ou +	
Masculino	-	5	11	8	=24
Feminino	-	6	4	1	=11
Total de jornalistas					=35

Como se pode constatar, não houve jornalistas na faixa que vai dos 7 aos 14 anos, que é algo óbvio, tendo em vista que até os 14 anos o indivíduo ainda não começou a fazer um curso superior.

Em relação ao grupo *nível de escolarização*, deve-se observar que nele só houve dados distribuídos em diferentes níveis na amostra da língua falada, uma vez que os dados da língua escrita foram todos produzidos por usuários com curso de nível superior em jornalismo. Isso se deu, logicamente, em função do gênero textual selecionado, que foram reportagens com foco no jornalismo científico.

Para a inserção dos dados no programa Varbrul, foi necessário codificá-los, agrupando-os de acordo com os grupos de fatores discriminados. O ideal é que essa codificação, por meio da atribuição de símbolos, tenha alguma relação com o fator que representa, a fim de que o linguista automatize a leitura dos dados no programa e possa fazer seu trabalho de forma mais precisa. Toda a codificação atribuída a cada fator dos grupos acima especificados está no Anexo 1 deste trabalho.

O próximo capítulo trará a apresentação dos resultados e as análises feitas a partir deles: em caráter quantitativo, com base nas rodadas feitas no Varbrul; e em caráter qualitativo, a partir das reflexões linguísticas e dos pressupostos teóricos do funcionalismo norte-americano.

CAPÍTULO 5

APRESENTAÇÃO E ANÁLISE DOS RESULTADOS

Após a identificação e seleção dos dados que interessam a esta pesquisa, feitas nos *corpora* especificados no capítulo anterior, obteve-se, em termos de frequência absoluta, um total de 302 dados na amostra da fala e 282 na da escrita.

Num primeiro momento, o interesse era saber a quantidade de finais desenvolvidas e reduzidas encontradas no conjunto de dados levantados – em outras palavras, o que se queria confirmar era se o número da variante reduzida era realmente superior ao da variante desenvolvida, conforme a hipótese previamente levantada. Assim, depois de os dados serem inseridos e codificados no programa computacional GOLDVARB X (SANKOFF; SMITH; TAGLIAMONTE, 2005) – conforme especificações já apontadas – o resultado obtido foi o especificado na Tabela 1:

TABELA 1

Quantificação das finais desenvolvidas e reduzidas nos dados da fala e da escrita

Amostra	Desenvolvidas	Reduzidas
<i>Fala</i>	9/302=3%	293/302=97%
<i>Escrita</i>	45/282=16%	237/282=84%

Fonte: PORTVIX , 2002; Superinteressante 2003, 2004.

Como se pode perceber, das 302 ocorrências da fala, 293 foram de orações reduzidas (97% do total) e apenas 9 de desenvolvidas (3% do total); na escrita, dos 282 dados, 237 corresponderam a casos de redução (84% do total), e 45 são exemplos de finais desenvolvidas (16% do total). Dessa forma, a quantificação das amostras da fala e da escrita das adverbiais finais é semelhante no que diz respeito ao tipo que mais ocorre, que é a reduzida. Na escrita, as finais desenvolvidas apresentam um percentual cerca de cinco vezes superior ao da fala,

enquanto nesta as reduzidas beiram a categoricidade. Assim sendo, a hipótese levantada na seção 1.2 se confirma, uma vez que era esperado que as finais reduzidas fossem realmente as mais frequentes nas duas bases de dados, assim como, apenas sendo comparadas as desenvolvidas, esperava-se que elas ocorressem em maior número na escrita que na fala. Essa maior frequência na escrita, acredita-se, se deve à predominância de estruturas impessoais num contexto discursivo mais formal, uma vez que se trata de um gênero textual que é controlado por padrões normativos da língua escrita.

Apenas para simples efeito comparativo, destacam-se na próxima tabela os resultados encontrados no trabalho de Finck (2000), ao levantar dados de orações adverbiais finais na fala da capital catarinense, ladeados pelos encontrados nesta pesquisa nos dados da fala da capital capixaba (PORTVIX):

TABELA 2

**Quantificação das finais desenvolvidas e reduzidas nos dados da fala:
Vitória (PORTVIX) versus Florianópolis (VARSL)**

Capitais	Desenvolvidas	Reduzidas
<i>Vitória</i>	9/302=3%	293/302=97%
<i>Florianópolis</i>	5/187=3%	182/187=97%

Fonte: PORTVIX, 2002; FINCK, 2000, p. 58.

Pela Tabela 2, vê-se que as duas capitais apresentaram percentuais de ocorrências para as variantes em análise exatamente iguais – 3% de desenvolvidas e 97% de reduzidas –, embora o total de dados seja diferente, isto é, 302 para Vitória e 187 para Florianópolis.

A decisão tomada por Finck (2000) em seu trabalho, tendo em vista a baixa ocorrência de dados para as desenvolvidas, foi a de não aplicar às orações finais um tratamento variacionista, mas analisá-las sob a perspectiva da gramaticalização²². Essa não é, no entanto, a decisão tomada nesta dissertação, uma vez que aqui são feitos alinhamentos e comparações do fenômeno na fala e na escrita, daí a manutenção do tratamento variacionista para a análise

²² Vale ressaltar que a pesquisadora também trabalhou com outros dois tipos de orações adverbiais, as temporais e as causais.

dos dados do PORTVIX, apesar da baixa frequência da variante final desenvolvida nos dados da fala.

Em relação a pesquisas sobre orações finais com base em dados da escrita, há o trabalho de Azevedo (2000), cuja base de dados remete à variante padrão do português escrito. O autor trabalhou com o rótulo “expressão da finalidade”, abarcando, além das orações adverbiais finais, as estruturas nominais com o traço semântico de finalidade. Assim, sendo excluídas essas estruturas, apresenta-se na Tabela 3 a comparação entre os resultados obtidos por este trabalho e os de Azevedo (2000, p.45):

TABELA 3
Quantificação das finais desenvolvidas e reduzidas nos dados da escrita:
Deoclécio (2011) versus Azevedo (2000)

Amostras	Desenvolvidas	Reduzidas
<i>Deoclécio</i>	45/282=16%	237/282=84%
<i>Azevedo</i>	157/915=17,2%	758/915=82,8%

Fonte: Superinteressante, 2003, 2004; AZEVEDO, 2000, p. 45.

Assim como aconteceu na comparação com os dados da fala, nesta, com os dados da escrita, os resultados para finais desenvolvidas e reduzidas são bastante aproximados em termos percentuais, apesar da grande diferença em termos absolutos. Assim, os números aferidos pelas comparações na fala e na escrita confirmam que o uso das orações finais como desenvolvidas ou reduzidas é um fenômeno da sintaxe da língua portuguesa, interno ao sistema, e não particular de uma ou outra região do Brasil, com as reduzidas sempre mais frequentes, isto é, como a variante não marcada.

Feitos esses comentários e observações gerais, é importante dizer que o desafio central deste estudo é buscar a compreensão de como se organiza a variação de ocorrências das finais desenvolvidas e reduzidas, que é maior nos dados da escrita que nos da fala, conforme projetam os números apresentados. Para essa empreitada, a análise e a ponderação de fatores internos e externos ao sistema linguístico são fundamentais. É o que se passa a fazer, portanto, nas próximas subseções: primeiro, apenas em termos percentuais e de frequência global dos

dados e, posteriormente, em termos de pesos relativos para as variáveis selecionadas pelo programa.²³

5.1 Resultados dos grupos de fatores de natureza linguística

5.1.1 Correferencialidade do sujeito da adverbial final

A primeira variável independente definida para a análise da variação sintática entre adverbiais finais desenvolvidas e reduzidas foi a *correferencialidade do sujeito da adverbial final*. Dentro desse grupo, fez-se uma subcategorização de acordo com a correlação ou não entre o sujeito da oração principal, ou da oração a que a final estivesse vinculada, e o sujeito desta. Nesse sentido, a Tabela 4 traz os resultados das ocorrências de acordo com o especificado na seção 5.2, quando se descreveu a subdivisão deste grupo de fatores em (a) *sujeito correferente ou igual*, (b) *sujeito parcialmente correferente* e (c) *sujeito não correferente ou diferente*. Nela, estão discriminadas a frequência absoluta – a quantidade de ocorrências – e a frequência relativa – os percentuais – dos dados referentes à amostra da fala e da escrita.

TABELA 4

Frequências absoluta e relativa dos dados da fala e da escrita, com base no grupo de fatores *correferencialidade do sujeito da adverbial final*

Sujeito	Fala		Escrita	
	Desenvolvidas	Reduzidas	Desenvolvidas	Reduzidas
(a) <i>correferente</i>	1/212=0.5%	211/212=99.5%	4/191=2.1%	187/191=97.9%
(b) <i>parc. correferente</i>	1/36=2.8%	35/36=97.2%	8/27=29.6%	19/27=70.4%
(c) <i>não correferente</i>	7/54=13%	47/54=87%	33/64=51.6%	31/64=48.4%
Total	9/302=3%	293/302=97%	45/282=16%	237/282=84%

Fonte: PORTVIX, 2002; Superinteressante, 2003, 2004.

²³ Ao longo das seções 5.1 e 5.2 são apresentados alguns pesos relativos para as variáveis independentes que não foram selecionadas pelo Varbrul como significantes do ponto de vista estatístico para a compreensão do fenômeno de variação linguística deste trabalho.

Os dados indicados nas tabelas desta seção devem ser assim lidos: na primeira coluna, estão discriminados os fatores definidos para o grupo; nas colunas seguintes, encontram-se os valores referentes à frequência absoluta – o número real de ocorrências da variante de acordo com o fator especificado – sobre o total entre as ocorrências das duas variantes, além da frequência relativa de cada fator. Por exemplo, olhando-se a segunda linha da coluna “Desenvolvidas”, nos resultados referentes à escrita, deve-se entender que há 8 dados dessa variante quando o fator em questão é *sujeito parcialmente correferente*. O número 27 se refere à soma das ocorrências das duas variantes quando esse fator está presente. Nesse sentido, a porcentagem 29.6% representa a frequência relativa, ou seja, das 27 ocorrências com sujeito parcialmente correferente, 29.6% delas estavam em finais desenvolvidas na amostra da escrita, o que implica dizer que sujeito parcialmente correferente eleva um pouco as desenvolvidas em relação à sua média, de 16% para 29.6%.

Como se pode observar nas frequências da Tabela 4, entre as ocorrências das finais desenvolvidas e reduzidas, quase houve representação de efeito categórico na variável independente especificada. Em outras palavras: houve 1 (uma) ocorrência apenas para finais desenvolvidas com sujeito correferente e 1 (uma) com sujeito parcialmente correferente nos dados da fala, o que resultou um percentual muito próximo de 100% dos casos para a variante final reduzida na presença desses dois fatores.

A partir de agora, serão comentados os resultados referentes a cada fator desta variável independente.

(a) Em relação ao fator *sujeito correferente*, constata-se que ele é favorecedor da ocorrência de finais reduzidas, pois aumenta a média dessa variante de 97% para 99.5% na fala, e de 84% para 97.5% na escrita. Por sinal, essa configuração sintática é a mais recorrente num período composto por subordinação cuja oração subordinada é uma adverbial final. Essa disposição linguística está diretamente vinculada à estrutura prototípica (GIVÓN, 1995, p. 29) do fenômeno aqui analisado, que é aquela em que há um sujeito na oração principal controlador do que se expressa na final (NEVES, 2000, p. 888), de maneira que a correferencialidade entre os seus sujeitos encaminha a oração final para a redução de infinitivo.

Mais adiante, na seção 5.4, após a discussão dos grupos de fatores selecionados por sua relevância estatística, este grupo de fatores será correlacionado a outros de natureza linguística para uma compreensão mais integrada da análise do fenômeno que aqui é

investigado. Por enquanto, serão examinados alguns exemplos, para que fique mais claro o favorecimento do fator em discussão. Também, nessas primeiras análises, após a colocação de um exemplo que ilustra a presença do fator em questão, é apresentada uma variante desse exemplo, passível de ocorrência na língua. Isso é feito com vistas a demonstrar que o fenômeno *finais desenvolvidas/reduzidas* , embora tenha se apresentado mais variável na escrita que na fala, é consistente e traz aspectos interessantes para uma análise linguística.

Primeiro, analise-se um dado do *corpus* da fala:

[...] mas procurei fazer o melhor dentro do curso PRA CONSEGUIR ABSORVER TUDO o que eu pudesse de melhor ali.

Neste dado, o sujeito da primeira oração, “eu”, é o mesmo da adverbial final. Ele age – *procurei fazer o melhor dentro do curso* – e aponta uma finalidade para si próprio – *conseguir absorver tudo o que eu pudesse de melhor ali* . Mesmo com baixíssima frequência nos dados codificados nesta pesquisa, uma possível variante desse exemplo é a paráfrase: *mas procurei fazer o melhor dentro do curso PRA QUE (EU) CONSEGUISSE ABSORVER TUDO o que eu pudesse de melhor ali* . O sujeito “eu”, posto entre parênteses, se justifica pelo fato de que poderia ou não ocorrer a sua explicitação. Estando expresso, traria certa funcionalidade discursiva, para marcar a ênfase de que a própria pessoa seria a beneficiada pelo que é declarado anteriormente. Entre as 19 entrevistas rastreadas no PORTVIX, o único dado real de final desenvolvida com sujeito correferente foi o seguinte: [...] *o ensino público, ele deve ser melhorado desde a base PRA QUE SEJA TÃO BOM QUANTO AS ESCOLAS PARTICULARES [...]* , em que “ele” – o ensino público – é o sujeito da oração núcleo e é repetido sem ser explicitado na oração final desenvolvida. De acordo com os resultados que foram obtidos, probabilisticamente as chances de esse dado ocorrer sob a forma de redução de infinitivo são muito maiores²⁴.

Seguindo com as exemplificações, examine-se um dado da escrita:

²⁴ Como se verá mais adiante, este grupo de fatores não foi selecionado como influente, do ponto de vista estatístico, no fenômeno variável deste trabalho, com base nos dados da fala. No entanto, na rodada 28, a primeira do *stepping down*, o fator apresenta 0.66 de peso relativo em relação à final reduzida, o que serve para indicar o seu favorecimento a essa variante, em uma eventual significância estatística da variável correferencialidade.

A cada mês, surge uma nova moda, uma nova cor, uma nova plástica, um novo penteado, um novo orgasmo que toda mulher tem que seguir PARA PODER EXERCER SEU PAPEL DE “PROTAGONISTA”.

Neste exemplo, embora não esteja clara a intenção do sujeito “toda mulher” em seguir os novos paradigmas sociais e de beleza de que trata a reportagem – pois o texto é escrito em terceira pessoa –, verifica-se que o sujeito da oração final, que vai *poder exercer seu papel de “protagonista”*, é também “toda mulher”. Nesse caso, o desenvolvimento da final é factível, sobretudo porque não é uma estrutura que codifica a intenção declarada prontamente pelo sujeito: [...] *um novo orgasmo que toda mulher tem que seguir PARA QUE POSSA EXERCER SEU PAPEL DE “PROTAGONISTA”*. Um exemplo real que comprova a possibilidade de ocorrência de final desenvolvida com uma estrutura semelhante a essa é: *Galinhas poedeiras vivem espremidas sob luz quase ininterrupta PARA QUE COMAM E BOTEEM OVOS SEM PARAR*, em que “galinhas poedeiras” é sujeito tanto da oração principal quanto da adverbial final.

(b) No que tange à correferencialidade parcial de sujeitos – quando algum termo da oração principal ou imediatamente anterior à final é nesta recuperado –, fazendo-se uma comparação da frequência em relação à média global das ocorrências, como é verificável na Tabela 4, acontece o seguinte: na amostra da fala, esse fator não aponta forte favorecimento nem para a final desenvolvida nem para a reduzida, pois a média de 3% da primeira vai para 2.8% na presença desse fator, enquanto a média de reduzidas, que é de 97%, vai para 97.2%. Por outro lado, nos dados da escrita, o comportamento é diferente, pois a ocorrência de *sujeito parcialmente correferente* na oração final tende claramente a favorecer o seu desenvolvimento, já que a média de 16% sobe para 29.6% na presença desse fator. Eis um exemplo da escrita:

O rei concedia terra aos melhores guerreiros PARA QUE PUDESSEM SE SUSTENTAR e, em troca, os cavaleiros estavam sempre preparados para lutar em nome do soberano.

Neste exemplo, cujo contexto discursivo remete à história do pagamento de impostos no Império Romano, o sujeito da oração principal é “o rei”, e o da final é “os melhores guerreiros”. Este é deduzido a partir de sua retomada, sem explicitação, na oração anterior, quando funciona como objeto indireto do verbo “conceder”, o que leva à correferência parcial aqui abordada. Uma paráfrase desse exemplo, com a variante reduzida, é possível no português brasileiro: *O rei concedia terra aos melhores guerreiros PARA PODEREM SE SUSTENTAR [...]*. De outra forma, um exemplo real da variante final reduzida, paralelo ao da

paráfrase feita, foi encontrado no *corpus* da escrita, com as mesmas características da desenvolvida acima especificada: [...] *o jornal mais antigo em circulação no Brasil, o Diário de Pernambuco, está por completar seu centenário, em 1925, e contrata o jovem sociólogo PARA ORGANIZAR UM LIVRO ALUSIVO À DATA*. Neste exemplo, pertencente a um texto que relata a trajetória pessoal e profissional do sociólogo Gilberto Freyre, o sujeito da oração final é “o jovem sociólogo”, que na oração anterior funciona como objeto direto do verbo “contratar”.

Agora, analise-se um exemplo da fala:

infelizmente ... o que seria mais certo seria eles ... tentarem amenizar:: a situação ...
fazendo ... a cabeça das crianças ... preparando... a cabeça das crianças ... PRA:: ...
RECEBER... UMA:: CERTA NOTÍCIA OU DE DIVÓRCIO... ou ... preparando pra ... mostrar
que:: o pai e a mãe dele que eles não estão bem

Para o entendimento dessa configuração sintática, deve-se observar que o sujeito da final é “as crianças”, recuperado da oração anterior, na qual funciona como adjunto adnominal. O contexto discursivo remete a uma situação em que os pais deveriam preparar as crianças para que estas recebessem a notícia de que eles estariam se divorciando. Nesse sentido, existe uma correferência parcial porque um termo da oração anterior, que não é o sujeito, é retomado como sujeito da oração final. É interessante observar que a variante final desenvolvida seria pertinente nesse mesmo contexto. Uma paráfrase desse exemplo pode ser: [...] *fazendo a cabeça das crianças, preparando a cabeça das crianças PRA QUE RECEBESSEM UMA CERTA NOTÍCIA OU DE DIVÓRCIO*. Ou, ainda, com o sujeito explicitado: [...] *fazendo a cabeça das crianças, preparando a cabeça das crianças PRA QUE ELAS RECEBESSEM UMA CERTA NOTÍCIA OU DE DIVÓRCIO*. No entanto, em circunstâncias reais, apenas foi encontrado 1 (um) dado de final desenvolvida com esse traço na amostra da fala: [...] *a educadora não aceitou o meu argumento... PRA QUE MEU DESENHO FICASSE NO PADRÃO DE TODAS AS PESSOAS me fez fazer um... refazer o desenho*. Nesse exemplo, o sujeito da final desenvolvida, “meu desenho”, é recuperado na oração seguinte, na qual “o desenho” funciona como objeto direto de “refazer”.

(c) Na Tabela 4, também é possível perceber que, quando o sujeito da oração final é diferente do da principal (não correferente), a alternância entre as variantes – final desenvolvida / final reduzida – ocorre de maneira mais perceptível, tanto na fala quanto na escrita. De acordo com os dados, esse fator aponta um aumento importante na média de finais desenvolvidas: na fala, o percentual é de 13%, o que significa um aumento de 10 pontos percentuais; na escrita, o

aumento é ainda mais significativo, pois a média de 16% sobe para 51.6%. Esse tipo de dado corresponde ao seguinte exemplo da língua falada:

[...] então PRA QUE ALGUMAS PESSOAS NÃO SE... TENHAM ALGUM PRIVILÉGIO IGUAL [...] eu acho que isso tem que ser avaliado de uma forma melhor

O informante, neste exemplo, fala a respeito de políticas públicas voltadas para a área da educação. O sujeito da final desenvolvida, que está anteposta à oração a que se vincula, é “algumas pessoas”. Esse termo não tem correferente na outra oração, cujo sujeito do verbo de opinião “achar” é o pronome “eu”. Sua forma reduzida poderia assim ocorrer: *então PRA ALGUMAS PESSOAS NÃO TEREM ALGUM PRIVILÉGIO IGUAL [...]eu acho que [...]*. Na amostra PORTVIX, foram encontrados 49 dados de final reduzida construída com sujeito diferente do da oração núcleo. Um desses dados é o seguinte: *então, eu acho, assim, que Vitória tem que ser uma/ uma cidade bem, bem bonita PRA NÃO/ AS PESSOAS SAÍREM DAQUI FALANDO BEM*. Como se observa, “as pessoas”, sujeito da oração final, não é igual ao das orações anteriores e não é retomado como termo de outra função sintática.

Em relação à escrita, segue um exemplo:

As terras seriam passadas para os filhos desde que houvesse a certeza de que eles tinham os mesmos genes que o pai. E, na ausência de testes de DNA, o jeito foi instituir a monogamia e manter a posse sobre a esposa PARA QUE OS PATRIARCAS FICASSEM SEGUROS QUANTO À SUA DESCENDÊNCIA.

Ao informar sobre elementos da biologia vinculados à ideia da monogamia, o exemplo da escrita apresenta uma oração final desenvolvida cujo sujeito “os patriarcas” não tem correspondente anafórico na oração anterior, que, por sinal, apresenta uma estrutura mais impessoal, no sentido de que não tem um sujeito com traço semântico [+humano]. Seria também possível a sua redução, como acontece em outro exemplo do *corpus* da escrita que apresenta estrutura idêntica: *No zoológico, tratadores escondem a comida pela jaula, PARA O ANIMAL PASSAR MAIS TEMPO PROCURANDO POR ELA, como na natureza*. Neste exemplo, verifica-se que a distinção dos sujeitos se confirma: “tratadores” é o da oração principal, e “o animal” o da adverbial final.

A relação entre a redução/desenvolvimento das orações finais e a correferencialidade de sujeitos será mais discutida na seção 5.4, quando se fizer o cruzamento dos resultados desta

variável com os da explicitação do sujeito da adverbial final, a fim de se perceber como a sua convergência interfere no fenômeno variável.

5.1.2 Características sintático-semânticas do sujeito da oração principal

A segunda variável independente escolhida para esta análise foi *características sintático-semânticas do sujeito da oração principal*²⁵. Como se viu na seção 4.2, este grupo de fatores foi considerado por se tratar de um traço recorrente da estrutura prototípica que contém a oração final, que é apresentar um sujeito controlador na oração principal que desencadeia o evento de finalidade expresso na adverbial final (NEVES, 2000, p. 888). Mesmo assim, foi verificado que existe certa variabilidade nesse tipo de característica, já que nem todos os sujeitos encontrados nos dados apresentam esse traço. Num primeiro momento, não se acreditava que esse grupo pudesse ter algum tipo de incidência significativa na variação entre finais desenvolvidas x reduzidas. Contudo, será visto mais adiante que, em pelo menos uma das modalidades analisadas, sua relevância estatística foi apontada após as rodadas de pesos relativos. Não sendo essa a principal intenção, é importante dizer que o objetivo de codificar esse tipo de característica estava muito mais atrelado ao fato de o pesquisador querer identificar o tipo de estrutura mais e menos recorrente no que diz respeito ao traço sintático-semântico do sujeito da oração principal, para uma posterior descrição que vai das estruturas mais marcadas às menos marcadas em cada ambiência discursiva. Feitas essas observações, observem-se as frequências para esta variável independente, na Tabela 5:

²⁵ Quando se fala “oração principal”, está sendo feita uma generalização, porque, na prática, muitas das orações às quais a final está vinculada mais proximamente podem não ser a principal, mas sim outra subordinada, por exemplo.

TABELA 5

Frequências absoluta e relativa dos dados da fala e da escrita, com base no grupo de fatores características sintático-semânticas do sujeito da oração principal

Características	Fala		Escrita	
	Desenvolvidas	Reduzidas	Desenvolvidas	Reduzidas
(a) <i>sujeito control./animado</i>	3/276=1.1%	274/276=98.9%	18/179=10.1%	161/179=89.9%
(b) <i>sujeito não control./inanimado/oracional</i>	6/26=24%	19/26=76%	27/103=26.2%	76/103=73.8%
Total	9/302=3%	293/302=97%	45/282=16%	237/282=84%

Fonte: PORTVIX 2002; Superinteressante, 2003, 2004.

De forma pragmática, o quadro que se desenha a partir dos números acima é o seguinte: em ambos os *corpora*, os fatores favorecedores das variantes em questão são os mesmos, uma vez que o fator (a), *sujeito controlador*, aumenta de forma pouco acentuada a média das finais reduzidas na fala – apenas 1.9 ponto percentual – e aumenta um pouco mais a média dessa mesma variante na escrita, projetando-a de 84% para 89.9%. Ao mesmo tempo, fica claro pelos resultados aferidos que o fator (b), *sujeito não controlador/inanimado/oracional*, é bastante favorecedor para a ocorrência das finais desenvolvidas nas duas bases de dados, já que na fala a média salta de 3% para 24%, e na escrita o desvio em relação à média é de pouco mais de 10 pontos percentuais.

Para melhorar o entendimento, vejamos alguns exemplos. Primeiro, um dado da fala:

meu pai já não podia... mais trabalhar... minha mãe... trabalhando direto... então eu tinha que:... estudar e:... logo arrumar alguma coisa PRA:: ... GANHAR DINHEIRO... eu ajudava meu pai...

Neste exemplo, quando o informante relata a necessidade que tinha de conseguir dinheiro para poder ajudar nas despesas da família, emerge na oração principal o sujeito “eu”, que estudava e precisava de um trabalho para poder ganhar dinheiro. O que se vê, portanto, é o sujeito da principal como controlador/agente do evento codificado na oração final. Como mostram os números, esse fator favorece a redução da oração final. Porém, fato é que há uma confluência de fatores, porque no próprio exemplo são encontradas três propriedades que

favorecem a redução oracional: sujeitos correferentes (conforme discorrido na seção 5.1.1), sujeito da oração principal com traço semântico controlador/agente e sujeito não explícito na oração final.

Mesmo assim, a amostra forneceu alguns exemplos – apenas 3 – de finais desenvolvidas cujo sujeito da oração principal tem o traço “controlador”, dos quais um é agora especificado: *eu tenho que gerenciar e organizar ... fazer um esquema ... PRA QUE TUDO FUNCIONE MUITO BEM*. Deve ser observado que há um sujeito na oração núcleo – “eu” – que expõe a finalidade da ação declarada, que é “gerenciar e organizar para tudo funcionar”.

Da base de dados da escrita, o seguinte exemplo foi pinçado:

PARA QUE VOCÊ SINTA O CHEIRO DE ALGO, é preciso que algum pedacinho desse objeto entre em contato com seu nariz.

De uma reportagem que fala sobre o sentido do olfato e a importância que a ciência moderna tem dado ao significado dos cheiros, foi extraído este exemplo. Observe-se que a estrutura que a oração final acompanha – *é preciso que algum pedacinho desse objeto entre em contato com seu nariz* – é formada de uma oração principal – *é preciso* – e de um sujeito oracional, a oração subordinada substantiva subjetiva *que algum pedacinho desse objeto entre em contato com seu nariz*. Nessa linha de análise, este tipo de estrutura foi codificado como sujeito não controlador, que, como se viu, é favorecedor da ocorrência de finais desenvolvidas. No entanto, assim como o exemplo da fala, este também apresenta uma confluência de fatores que favorecem o desenvolvimento da oração final: os sujeitos da principal e da final não são correferentes; o sujeito da principal não é controlador; o sujeito da final é explícito. A explicitação ou não dos sujeitos nas finais será abordada na seção 5.1.4.

Da mesma forma que se procedeu anteriormente com o exemplo da fala, para o da escrita também há dados que permitem confirmar a possibilidade de ocorrência da variante final reduzida na presença do fator *sujeito não controlador/inanimado/oracional*. Foram 76 dados com essa propriedade, entre os quais se destaca o seguinte exemplo: *O clima do pós-guerra é de festa animada pelo jazz. O corte de cabelo é curto PARA EQUILIBRAR PEQUENOS CHAPÉUS*. É possível identificar na primeira oração que o tipo de sujeito expresso – *o corte de cabelo* – não tem o traço de controlador da ação seguinte, nem tampouco de animação. Entretanto, na sequência, a oração final apresenta-se reduzida de infinitivo.

Não se pode deixar de notar uma característica importante: tendo ocorrido a final reduzida com sujeito do tipo não controlador, como acaba de ser visto, esses casos têm uma marca bastante recorrente, que é a da não expressão do sujeito na oração final. Dos 76 casos, apenas 3 apresentaram sujeito explícito, como em:

A vitela é a carne de um bezerro anêmico que passa seus cinco meses de vida em um cercado minúsculo, impedido de se mover, PARA A CARNE FICAR MACIA.

5.1.3 Voz/aspecto semântico do verbo da adverbial final

O grupo de fatores *voz/aspecto semântico do verbo da adverbial final* foi escolhido para compor esta análise em função de uma percepção que se teve ao serem observados os primeiros dados dos *corpora* que compõem esta pesquisa: quando a oração final estava na voz passiva, ela tendia a desenvolver-se. Ao mesmo tempo, era preciso controlar como se davam as ocorrências na voz ativa, que são os casos mais frequentes e pareciam estar mais vinculados ao uso de finais reduzidas. Como em alguns casos o verbo se apresenta na voz ativa, sem que o sujeito tenha traços de agentividade, incorporou-se ao fator voz o traço semântico do verbo da oração final, de modo a indicar o sujeito como agente ou paciente na expressão de finalidade. Para analisar os resultados encontrados, verifiquem-se as frequências na Tabela 6:

TABELA 6

Frequências absoluta e relativa dos dados da fala e escrita, com base no grupo de fatores *voz/aspecto semântico do verbo da adverbial final*.

Voz/aspecto semântico do verbo	Fala		Escrita	
	Desenvolvidas	Reduzidas	Desenvolvidas	Reduzidas
(a) ativa /sujeito como agente	8/289=2.8%	281/289=97.2%	37/260=14.2%	223/260=85.8%
(b) passiva / sujeito como paciente	1/13=7.7%	12/13=92.3%	8/22=36.4%	14/22=63.6%
Total	9/302=3%	293/302=97%	45/282=16%	237/282=84%

Fonte: PORTVIX, 2002; Superinteressante, 2003, 2004.

Apesar da frequência massiva de verbos na voz ativa ou que expressam o sujeito da final como agente, constatada nas duas bases de dados – fato que pode justificar a não seleção estatística desta variável, como se verá adiante –, o que se esperava a respeito das tendências de favorecimento à ocorrência das variantes em estudo se confirmou em termos percentuais, porque tanto na fala quanto na escrita a média das finais desenvolvidas é aumentada quando o fator (b), *voz passiva / verbo que expresse o sujeito como paciente*, está presente: o aumento é de 3% para 7.7% na fala (peso relativo de 0.57), e de 16% para 36.4% na escrita (peso relativo de 0.65). Na própria tabela, por outro lado, vê-se que a média das finais reduzidas aumenta de forma pouco expressiva na presença do fator (a), apresentando pesos relativos de 0.50 e 0.51²⁶, na fala e na escrita, respectivamente. É importante frisar que os pesos relativos aqui apresentados são meramente ilustrativos, pois, como será visto na seção 5.3, os fatores desta variável não apresentaram resultados estatísticos consistentes para atestar a sua influência na variação das orações finais.

Para ilustrar a descrição dos dois fatores, seguem-se os exemplos:

1. **Dado da fala:** investindo ... pesado na sa/ na:: saúde ... na educação ... e na segurança ... primeiro ... o primordial pra mim seria a educação ... o primordial ... porque pessoas cultas não vão cair no [hábito] ... facilmente ... segundo ... saúde ... PRA:: ... QUE AS PESSOAS... TENHAM COMO... PROSEGUIR NESSA CAMINHADA...
2. **Dado da escrita:** No Japão, há empresas que borrifam aromas no sistema de ar-condicionado dos escritórios - brisa de limão pela manhã, lufadas de manjerição à tarde – PARA QUE SEUS FUNCIONÁRIOS TRABALHEM MAIS SATISFEITOS.
3. **Dado da fala:** [...] então acho que tem que ser... tem que ter uma forma melhor de/ de analisar de avaliar PRA QUE REALMENTE OS PRIVILÉGIOS DO GOVERNO SEJAM DADOS A QUEM NECESSITA.
4. **Dado da escrita:** ele criou condições PARA QUE POVOS NÃO-JUDAICOS, ao receberem a mensagem de Deus, FOSSEM INSERIDOS DE FORMA IGUALITÁRIA NA COMUNIDADE CRISTÃ.

Esses quatro exemplos são casos de finais desenvolvidas dos dois *corpora*, conforme destaque. Em (1) e (2), os verbos estão dispostos na voz ativa; já em (3) e (4) se dispõem na voz passiva analítica. Nos dois últimos exemplos, como se aferiu, está o fator que favorece esse tipo de variante. Contudo, vale assinalar que, mesmo com uma tendência probabilística

²⁶ Pesos relativos retirados da rodada 28 (fala) e da rodada 20 (escrita) do *stepping down*.

de ocorrência de uma final desenvolvida na presença do fator (b), só houve 1 (um) dado encontrado nas células do PORTVIX com essa característica, que é o exemplo 3, supracitado.

Agora, alguns exemplos de finais reduzidas:

5. **Dado da fala:** e a gente/ a gente senta se reúne PRA DISCUTIR AS QUESTÕES mas eu ... eu não concebo ser uma coordenadora só de ... assim ... obedecer ordens
6. **Dado da escrita:** PARA PASSAR AO ESTÁGIO DA ECONOMIA ECOLÓGICA, a humanidade teria de assumir os riscos de falência em massa, crises de governos e dos sistemas financeiros.
7. **Dado da fala:** aí no outro mês seguinte ‘cê vai ... tem greve ... a gente já ficou uns quatro meses assim ... PRA ... PRA SER ATENDIDA ...
8. **Dado da escrita:** PARA SER CONSIDERADO BONITO, um nariz deve ser virtualmente invisível, não pode chamar a atenção.

Nesses casos de finais reduzidas, a disposição é a mesma que se fez nos exemplos anteriores: em (5) e (6), os verbos estão dispostos na voz ativa; e, em (7) e (8), a voz é a passiva analítica. Observe-se que, em (5) e (6), estão os casos mais recorrentes nas duas bases de dados, inclusive em termos absolutos, pois foram 281 ocorrências no PORTVIX, e 223 nas reportagens da Superinteressante.

Apesar de se estar analisando neste momento a questão da voz do verbo, seria imprudente deixar de assinalar uma característica dos exemplos das finais desenvolvidas que se mostra crucial para este trabalho e que emerge dos quatro primeiros exemplos supracitados: todos os sujeitos são não correferentes e estão explícitos. De igual modo, observando-se os exemplos das finais reduzidas, traços opostos aos anteriores são detectados: todos os sujeitos são correferentes e não se explicitam. Esses cruzamentos serão discutidos mais adiante, após a definição das variáveis independentes com relevância estatística apontadas pelo Varbrul.

5.1.4 Explicitação do sujeito na adverbial final

Outro grupo de fatores definido para este estudo foi o da *explicitação do sujeito na adverbial final*. A escolha desse fator, de natureza linguística, foi motivada pela questão da gradiência na estrutura das orações finais, principalmente quando se pensa no princípio da *marcação* (GIVÓN, 1995, p. 28), em seu critério da *complexidade formal*. Como este trabalho busca mostrar que finais desenvolvidas e reduzidas não são formas meramente discretas, mas

gradientes, a explicitação ou não do seu sujeito contribui para a configuração desse tratamento escalar. No entanto, esta variável já foi timidamente explorada e correlacionada a outros grupos de fatores nas seções anteriores, o que indica que ela cumpre um papel de relevância na área de variação do fenômeno investigado por este estudo.

Para seguir com a análise, deve ser observada a Tabela 7 para a aferição das frequências deste grupo de fatores:

TABELA 7

Frequências absoluta e relativa dos dados da fala e da escrita, com base no grupo de fatores explicitação do sujeito da adverbial final

Sujeito da oração final	Fala		Escrita	
	Desenvolvidas	Reduzidas	Desenvolvidas	Reduzidas
(a) <i>explícito</i>	8/74=10.8%	66/74=89.2%	36/42=85.7%	6/42=14.3%
(b) <i>não explícito</i>	1/228=0.4%	227/228=99.6%	9/240=3.8%	231/240=96.2%
Total	9/302=3%	293/302=97%	45/282=16%	237/282=84%

Fonte: PORTVIX, 2002; Superinteressante, 2003, 2004.

Como se pode identificar nas frequências da Tabela 7, a presença do fator *sujeito explícito* na subordinada favorece de forma evidente a variante final desenvolvida, pois, como a média de desenvolvidas nos dados da fala é de 3% e nos da escrita é de 16%, essa média é aumentada para 10.8% e 85.7%, respectivamente, quando a explicitação do sujeito é verificada. O mesmo favorecimento acontece para as finais reduzidas quando o fator *sujeito não explícito* está presente: com 99.6% na fala e 96.2% na escrita, esse fator aumenta a média de redução das orações finais, que é de 97% e 84%, respectivamente.

Examinem-se, portanto, alguns exemplos para que se perceba melhor a configuração dos fatores agora em questão. Segue um dado do PORTVIX:

Não, a gente vai lá mais PRA VISITAR MESMO O LOCAL, encontrar, todo mundo vai, a gente vai pra pra se encontrar mesmo.

De acordo com este exemplo, as evidências antes descritas são visíveis: a informação de que um grupo de pessoas vai para um local para se reunir é codificada na final reduzida sem a

explicitação do sujeito. Como se sabe, a não explicitação do sujeito na oração final está intimamente vinculada à sua correferencialidade, uma vez que o fato de os sujeitos da principal e da final serem iguais – *a gente* – motiva, segundo o princípio da economia linguística (HAIMAN, 1983), a não realização de um dado discursivamente conhecido. Apesar de menos frequentes, houve no *corpus* 66 casos de finais reduzidas em que o sujeito foi expreso. Uma das ocorrências é a seguinte: [...] *mas aí PRA CÊ IR MORAR NUM SÍTIO, cê tem que produzir lá [...]*, em que o sujeito “cê” se explicita, mesmo sendo correferente ao da oração principal. Ainda assim, há marcas estruturais que diferenciam um pouco este exemplo do anterior, haja vista a posição da oração final não ser a mesma nos dois casos. Parece ser que a posição anteposta da oração final em relação à principal propicia a explicitação do sujeito naquela, que é o que ocorre no segundo exemplo.

Para confirmar as constatações anteriores, assinala-se um exemplo da base de dados da língua escrita:

Dizem também que o novo método deveria ser comparado com o melhor método disponível no país PARA QUE O GOVERNO LOCAL POSSA FAZER uma escolha entre o método em uso e o novo.

Neste dado, corrobora-se o que foi antes afirmado. Na oração final destacada, encontra-se expreso o sujeito *o governo local*, que não é correferente ao da oração principal, *o novo método*. Mesmo com menor incidência na amostra da escrita, com apenas 6 casos, segue-se um exemplo de final reduzida com sujeito explícito, para efeitos de confirmar a possibilidade de variação de uma ocorrência como a do exemplo acima: *Alguns levam o bichinho a um psicólogo de animais, mas, quando descobrem que eles também precisam mudar seus hábitos, acabam aceitando a sugestão de dar remédios como Prozac ou Valium PARA O MASCOTE FICAR MAIS TRANQUILO*. Desse modo, deve ser observado que o sujeito da oração final reduzida de infinitivo é *o mascote*, que não é correferente ao da oração anterior, o que favorece sobremaneira a sua expressão.

Assim como se falou em motivação por economia, também se pode falar em motivação por iconicidade, pois, juntando-se as características encontradas nos dois exemplos, resulta que a correferência e a não explicitação dos sujeitos motivam a ocorrência da final reduzida, ao passo que a não correferência e a explicitação dos sujeitos motivam a emersão discursiva da final desenvolvida. Ou seja, pelo princípio da economia, o que é menos previsível na língua é

mais codificado, expresso, enquanto o que é mais previsível é menos codificado, menos expresso; pela motivação icônica, o que ocorre neste fenômeno é que a menor codificação tende a fazer ocorrer uma forma menos extensa da variável dependente, que é a final reduzida. O mesmo ocorre com a final desenvolvida, que é mais extensa, motivada iconicamente por ter mais elementos codificados. Esta última constatação é perceptível no dado da final desenvolvida supracitado: a final é estruturada com PARA + QUE + SUBJUNTIVO, o que indica a forma típica da variante desenvolvida, cujo sujeito – *o governo local* – se explicita por uma necessidade discursiva, já que não é correferente nem ao da principal, que apresenta traço de indeterminação, nem ao da subordinada substantiva que a antecede, o sujeito paciente *o novo método*.

5.1.5 Posição da adverbial final

A quinta e última variável independente de natureza linguística proposta para a investigação do fenômeno de variação entre as finais desenvolvidas e reduzidas é a da *posição da adverbial final* em relação à oração principal, se posposta ou anteposta a esta, conforme se descreveu na seção 5.2²⁷. Os resultados encontrados para esse grupo de fatores estão expressos na Tabela 8:

TABELA 8

Frequências absoluta e relativa dos dados da fala e da escrita, com base no grupo de fatores *posição da adverbial final*

Posição	Fala		Escrita	
	Desenvolvidas	Reduzidas	Desenvolvidas	Reduzidas
<i>(a) final posposta</i>	7/281=2.5%	274/281=97.5%	36/206=17.5%	170/206=82.5%
<i>(b) final anteposta</i>	2/21=9.5%	19/21=90.5%	9/76=11.8%	67/76=88.2%
Total	9/302=3%	293/302=97%	45/282=16%	237/282=84%

Fonte: PORTVIX, 2002; Superinteressante, 2003, 2004.

²⁷ Outro tipo de posição estrutural possível das adverbiais finais é a intercalada. No entanto, não foram encontrados dados com essa propriedade nos *corpora* utilizados nesta pesquisa.

Apesar de não ter sido selecionada por sua relevância estatística para a interpretação variável das adverbiais finais, esta variável independente será brevemente analisada nesta seção a partir das frequências obtidas. Como se pode constatar, tanto nos dados da fala quanto nos da escrita, a posição da oração final em relação à principal é predominantemente posposta. Numa comparação horizontal entre os dois fatores do grupo, constata-se que há 93% (281/302) de posposição e 7% (21/302) de anteposição das finais na fala. Na escrita, a maior frequência se mantém, com 73% (206/282) de finais pospostas e 27% (76/302) de antepostas. Essa proporção é bastante parecida com os resultados encontrados por Finck (2000, p.66), que trabalhou exclusivamente com dados da fala de Florianópolis: nas desenvolvidas, com todos os casos sendo de finais pospostas, enquanto 94% das reduzidas se apresentaram colocadas após a oração principal. Em pesquisa com dados da língua inglesa, Thompson (1985) encontrou essa mesma predominância em termos da maior frequência de finais pospostas, o que leva à conclusão de que essa característica extrapola os limites da língua portuguesa, ou, ainda mais, das línguas neolatinas. Na verdade, há aí uma relação icônica, porque é prototípico na expressão da finalidade que se declare “o que se faz” primeiro para depois ser apontado “o fim”, “o objetivo” daquilo que se fez anteriormente.

Nos dos dados do PORTVIX, a anteposição da oração final tende a favorecer o seu desenvolvimento, porque a sua frequência relativa, de 9.5%, é maior que a média de finais desenvolvidas, de apenas 3% dos dados coletados. Em relação à posposição, os resultados mostram que este fator não tem um comportamento de favorecer ou desfavorecer de forma evidente uma ou outra variante, comparada sua frequência relativa com a média das ocorrências. Já com base nos dados da escrita, essa relação é um pouco diferente: a média de finais desenvolvidas, que é de 16%, aumenta em apenas 1.5 ponto percentual quando a oração é posposta; a das reduzidas, que é de 84%, aumenta 4.2 pontos percentuais na ocasião em que a final é anteposta.

Os resultados atestam, pois, que nas duas bases de dados a variável *posição* apresenta efeitos diferentes: a posposição, mesmo sendo considerado o discreto aumento de média que confere às reduzidas na fala e às desenvolvidas na escrita, não se apresenta como grande favorecedora ou desfavorecedora de uma ou de outra variante; a anteposição, por outro lado, tende a favorecer a ocorrência de desenvolvidas na fala e de reduzidas na escrita.

A primeira tentativa na busca por uma resposta para essa inversão nos dois *corpora* resultou no cruzamento deste grupo de fatores com outro de caráter linguístico, o da *explicitação do*

sujeito da oração final. De acordo com os efeitos de favorecimento às variantes da variável dependente anteriormente descritos, o quadro é o seguinte: as duas ocorrências de finais desenvolvidas antepostas na fala contêm sujeitos explícitos, o que atende às expectativas para a estrutura dessa variante. São estes os dois casos:

- (1) [...] então PRA QUE ALGUMAS PESSOAS NÃO SE... TENHAM ALGUM PRIVILÉGIO IGUAL
[...]eu acho que isso tem que ser avaliado de uma forma melhor
- (2) [...] a educadora não aceitou o meu argumento... PRA QUE MEU DESENHO FICASSE NO
PADRÃO DE TODAS AS PESSOAS me fez fazer um... refazer o desenho

No entanto, esse entrelaçamento não é eficaz quando se pretende justificar a posição da oração em função da explicitação do sujeito, pois a relação de dependência parece ser contrária: a explicitação é que depende da posição²⁸. Tanto é assim que na apuração dos dados, dos 67 casos de reduzidas antepostas na escrita, apenas 3 apresentaram sujeito explícito. Fato que corrobora o que já se apurou neste trabalho – a explicitação do sujeito favorece as finais desenvolvidas, e a não explicitação, as reduzidas.

Para seguir com a análise, seguem alguns dados do *corpus* da escrita:

- (3) PARA SE OBTER CHEIRO DE BAUNILHA, ninguém mais precisa das favas colhidas em Madagáscar: produz-se vanilina em laboratório.
- (4) PARA SE ALCANÇAR ESSE ESTÁGIO, os sistemas tributários mundiais precisariam ser reformulados de modo a oferecer subsídios à reciclagem e à geração de energia limpa e renovável e taxar atividades insustentáveis, como o uso de combustível fóssil.
- (5) PARA O FIEL VIRAR PASTOR, o importante é ter carisma.

Consoante se verifica, esses três exemplos têm uma configuração linguística bem diferente da encontrada nos dados da fala, anteriormente expostos. A redução nestas três ocorrências está acompanhada de características que lhe são pouco comuns, já que as orações finais, mesmo reduzidas, se apresentam aqui com sujeitos expressos e não correferenciais, o que é motivado por sua anteposição, pois, sendo a oração final anteposta e o seu sujeito não correferente, espera-se que este esteja expresso, uma vez que se trataria de uma informação potencialmente nova no discurso. Outra característica importante de ser observada é que os exemplos (3) e

²⁸ Conferir uma breve explanação sobre isso na seção 5.6.

(4), por serem construções na voz passiva, apresentam sujeitos explícitos que não têm o mesmo traço sintático-semântico do que está em (5), o que também pode ter interferido no resultado encontrado.

Assim, a questão de o mesmo fator – *final anteposta* – favorecer variantes distintas nas diferentes bases de dados deve estar vinculada às peculiaridades dos gêneros discursivos dos quais eles foram extraídos– entrevista no PORTVIX, e reportagem na Superinteressante –, em virtude dos domínios funcionais que deles emergem. Nos termos de Biber (1988, p. 36), não há caracterização linguística que seja única e verdadeira em todos os gêneros falados e escritos, pois há muitas semelhanças e diferenças entre eles, em função de aspectos intra e extralinguísticos²⁹. Braga *et al* (2008, p. 203), ao abordarem a ordem das orações em contextos de fala e escrita, problematizam as razões para as disparidades encontradas na conformação dos dados numa comparação entre fala e escrita, e trazem para a discussão as ponderações de Biber (1988):

[...] o cotejo entre fala e escrita pressupõe rigor no que tange aos tipos e gêneros textuais nos quais são coletadas as ocorrências que fundamentam a análise. Uma vez que a distribuição de numerosas propriedades gramaticais é sensível aos gêneros textuais, a utilização de gêneros textuais diferentes, como base de dados, pode levar a resultados conflitantes.

No entanto, como se sabe, este trabalho utiliza dados da fala e da escrita de duas amostras compostas por exemplos de gêneros textuais diferentes. Também é sabido que, quando comparadas fala e escrita, a única variável independente de caráter linguístico atribuída à análise que apresentou tendências de favorecimento diferentes a cada uma das variantes – final desenvolvida e reduzida – foi a posição da oração final. Assim, mesmo as ponderações de Biber (1988) e Braga *et al* (2008) sendo de uma linha metodológica bastante coerente, a escolha dos diferentes gêneros textuais utilizados nesta dissertação, para que se fizesse uma análise comparativa entre fala e escrita, não pode ser invalidada, sobretudo pela consistência

²⁹ Paráfrase e adaptação livre a partir do seguinte trecho: [...] there is no linguistic or situational characterization of speech and writing that is true of all spoken and written genres. On the one hand, some spoken and written genres are very similar to one another (e.g., public speeches and written exposition). On the other hand, some spoken genres are quite different from one another (e.g. conversation and public speeches), as are contextual differences between typical speech and writing. These other components include social categories (e.g. ,class, ethnic group, sex and age), individual personality characteristics of the speaker and addressee, and the social role relationship (e.g., power and status) between speaker and addressee [...] (BIBER, 1988, p.36)

dos resultados encontrados no que diz respeito às similaridades e diferenças, que são passíveis de compreensão. Em suma, as diferenças quantitativas encontradas, para mais e para menos, em termos das frequências – absoluta e relativa – não apontaram nas outras variáveis independentes tendências de favorecimento divergentes para o fenômeno de variação linguística investigado.

Na próxima seção, serão discriminados e analisados os resultados encontrados a partir da codificação das variáveis independentes de caráter social: sexo/gênero; faixa etária e nível de escolarização.

5.2 Resultados dos grupos de fatores de natureza social

As variáveis independentes de ordem social selecionadas para esta análise foram definidas com base na amostra da fala – PORTVIX –, que delimitou a escolha dos informantes para compor o conjunto de suas entrevistas por meio das variáveis *sexo/gênero*, *faixa etária* e *nível de escolarização*. Em função disso, os dados encontrados no *corpus* da língua escrita foram codificados tendo em vista esses mesmos grupos de fatores. No entanto, por causa do gênero textual escolhido para a composição desse *corpus*, o *nível de escolarização* não foi atribuído de forma direta na codificação dos dados, já que todas as reportagens foram redigidas por profissionais com curso superior de jornalismo.

É importante dizer que se tomou por hipótese inicial que fatores sociais não apresentariam relevância estatística para o caso de variação sintática entre as finais desenvolvidas e reduzidas, porque, por evidências e recorrências, trata-se de um fenômeno governado mais fortemente por fatores internos do sistema linguístico. Apesar disso, será visto que a variável referente ao grau de escolaridade dos informantes do PORTVIX tem papel importante na área de variação do fenômeno em estudo, pois apresentou significância estatística, o que vai ao encontro do que se tomou por hipótese na formulação inicial desta pesquisa.

Nesse sentido, é válido fazer referência ao que se ponderou na seção das hipóteses deste trabalho (cf. seção 1.2): a variação sintática entre as finais desenvolvidas e reduzidas não apresenta marcas de estigma social e não causa, por exemplo, o mesmo impacto provocado por variantes no âmbito da concordância verbal e nominal que sejam desviantes da variante padrão da língua, como em “nós vai” ou “pega aqueles copo pequeno”.

Feitas essas observações, passa-se agora à aferição dos resultados encontrados para cada um dos três grupos de fatores especificados.

5.2.1 Sexo/gênero

O primeiro grupo de fatores de natureza social a ter seus números apresentados é o de *sexo/gênero*, conforme a Tabela 9:

TABELA 9

Frequências absoluta e relativa dos dados da fala e da escrita, com base no grupo de fatores *sexo/gênero*

Sexo/gênero	Fala		Escrita	
	Desenvolvidas	Reduzidas	Desenvolvidas	Reduzidas
<i>(a) masculino</i>	4/161=2.5%	157/161=97.5%	30/199=15.1%	169/199=84.9%
<i>(b) feminino</i>	5/141=3.5%	136/141=96.5%	15/83=18.1%	68/83=81.9%
Total	9/302=3%	293/302=97%	45/282=16%	237/282=84%

Fonte: PORTVIX, 2002; Superinteressante, 2003, 2004.

Em termos das frequências, os números mostram que homens e mulheres usam mais as finais reduzidas que as desenvolvidas para se expressar, tanto na fala quanto na escrita, pela razão simples de que há mais reduzidas que desenvolvidas no recorte das amostras pesquisadas. Comparativamente, verifica-se que as mulheres apresentaram um índice levemente superior ao dos homens no desenvolvimento das finais: com 3.5% na fala, aumentando em 0.5 ponto percentual a média global (peso relativo de 0.73), e 18.1% na escrita, aumentando em relação à média 2 pontos percentuais (peso relativo de 0.59). Nessa mesma linha de raciocínio, vê-se que o fator *masculino* atua, de forma discreta, como favorecedor das finais reduzidas: sobe a média de 97% para 97.5% na fala – com peso relativo de 0.70 – e de 84% para 84.9% na escrita – com peso relativo de 0.54.³⁰

³⁰ Os pesos relativos especificados foram retirados da rodada 28 (fala) e da rodada 20 (escrita) do *stepping down* e não apresentam significância estatística, sendo usados apenas para conjecturas.

Todavia, como será visto na seção 5.3, que tratará dos grupos de fatores selecionados como estatisticamente relevantes para a variação entre as finais desenvolvidas e reduzidas, a variável *sexo/gênero* não atua determinantemente para isso. Em geral, os estudos que consideram a variação linguística entre homens e mulheres têm uma dimensão voltada para questões de mudança linguística, como perceber indícios de formas de menos prestígio competindo com ou ocupando o espaço de formas mais prestigiadas socialmente.

Segundo Paiva (2008. p.35), muitos “[...] estudos sobre processos variáveis do português apontam para o que poderíamos denominar uma maior consciência feminina do *status* social das formas linguísticas”. Assim, as mulheres seriam, em alguns casos, um vetor de manutenção do padrão linguístico – dependendo da comunidade de fala a que pertençam –, o que revelaria certo conservadorismo em sua fala. Outros estudos consideram que as mulheres estariam na vanguarda da mudança linguística de determinados fenômenos, seja em aspectos fonético-fonológicos, morfossintáticos ou semânticos. É certo que essa configuração dependerá do tipo de mudança linguística e também da confluência do fator *sexo/gênero* com outros de cunho social, tais como a faixa etária, a classe social e o nível de escolarização do usuário da língua (LABOV, 2008).

Mas se encaixaria o fenômeno deste estudo nessas características? A resposta é não, sobretudo pelo fato de a relevância estatística da variável não ter se confirmado em nenhum dos dois *corpora*. São as finais desenvolvidas formas mais conservadoras que as reduzidas? Talvez não sejam mais conservadoras, mas sim ocorram com maior frequência numa ambiência discursiva mais formal, tanto na expressão de mulheres, como na de homens.³¹

No que diz respeito à modalidade escrita, deve-se considerar o fato de que a amostra aqui utilizada – o gênero reportagem – não pode receber o mesmo tratamento analítico da fala. Esse gênero apresenta um nível de linguagem que se coaduna com o modelo linguístico da gramática normativa, não representando a mesma espontaneidade das entrevistas do *corpus* da fala. Dessa maneira, levando-se em consideração que há certa formatação e organização textuais a serem seguidas dentro das redações de jornais e revistas, até que ponto se pode medir a diferença entre a reportagem escrita por uma mulher da escrita por um homem? Seria, portanto, um erro analítico fazer considerações muito aprofundadas a respeito da variável

³¹ Esse tipo de constatação fica como desafio para futuros trabalhos que comparem a variação do fenômeno em ambientes discursivos diferentes em seus graus de formalidade e informalidade, na modalidade falada e escrita da língua portuguesa.

sexo\gênero e suas possíveis influências nesta análise, já que não há evidências muito claras de como essa influência se daria.

5.2.2 Faixa etária

Ainda na esteira das variáveis sociais, estão discriminadas na Tabela 10 as frequências referentes ao grupo de fatores *faixa etária*, para os dois *corpora*:

TABELA 10

Frequências absoluta e relativa dos dados da fala e da escrita, com base no grupo de fatores *faixa etária*

Faixa etária	Fala		Escrita	
	Desenvolvidas	Reduzidas	Desenvolvidas	Reduzidas
(a) 7-14 anos	0/35=0%	35/35=100%	----	----
(b) 15-25 anos	4/119=3.4%	115/119=96.6%	12/80=15%	68/80=85%
(c) 26-49 anos	5/74=6.8%	69/74=93.2%	22/139=15.8%	117/139=84.2%
(d) 50 anos ou mais	0/74=0%	74/74=100%	11/63=17.5%	52/63=82.5%
Total	9/302=3%	293/302=97%	45/282=16%	237/282=84%

Fonte: PORTVIX, 2002; Superinteressante 2003, 2004.

Em relação ao que mostra a Tabela 10, pode-se ver que há efeito categórico (100% das ocorrências) em dois dos fatores referentes à amostra PORTVIX: todos os dados codificados na faixa que vai dos 7 aos 14 anos e na que vai dos 50 anos em diante são de finais reduzidas, o que evidenciaria um padrão curvilíneo, característico de variação estável na comunidade de fala e da mudança do indivíduo ao longo da vida. Também se verifica que a faixa etária que mais favorece as finais desenvolvidas é a que vai dos 26 aos 49 anos, pois com 6.8% de frequência relativa aumenta a média das desenvolvidas em 3.8 pontos percentuais.

Para que fosse viável fazer a rodada de pesos relativos, os dois fatores com 100% de ocorrências em apenas 1 (uma) das variantes em estudo foram amalgamados aos que mais próximos deles estavam na linha etária que divide cada um deles: o fator (a), *7-14 anos*, foi

amalgamado com o fator (b), *15-25 anos*, e o fator (d), *50 anos ou mais*, juntou-se ao fator (c), *26-49 anos*. Esse processo será mais bem explicado na seção 5.3.

Sobre os dados da escrita, primeiro devem ser explicadas as ausências de dados na faixa etária que vai dos 7 aos 14 anos: todos os jornalistas tinham mais de 23 anos quando escreveram as reportagens³², o que impossibilitou a ocorrência de dados entre as crianças e os adolescentes.

Uma leitura vertical dos resultados da escrita permite constatar que há um equilíbrio entre os três fatores: o que mais favorece a variante desenvolvida é o da faixa dos 50 anos ou mais, com 17.5% de frequência relativa. Mas essa frequência é apenas 1.7 ponto percentual superior à da faixa dos 26 aos 49 anos de idade. O mesmo acontece quando se projeta o favorecimento das reduzidas: com 85%, a faixa que vai dos 15 aos 25 anos é a que mais as favorece; no entanto, a fator (c) – faixa dos 26 aos 49 anos – só lhe é inferior em 0.8 ponto percentual. Ou seja, a distribuição das variantes da variável dependente nos dados da escrita é bem equilibrada quando o grupo de fatores em análise é a faixa etária de quem produziu os textos, o que não oferece um efeito percentual de grande relevância.

5.2.3 Nível de escolarização

O terceiro e último grupo de fatores sociais é o *nível de escolarização*. Os resultados obtidos tomando por base os três fatores especificados (níveis fundamental, médio e superior) podem ser vistos na Tabela 11:

TABELA 11

Frequências absoluta e relativa dos dados da fala, com base no grupo de fatores *nível de escolarização*

Nível de escolarização	Desenvolvidas	Reduzidas
<i>(a) fundamental</i>	0/90=0%	90/90=100%
<i>(b) médio</i>	1/93=1.1%	92/93=98.9%
<i>(c) superior</i>	8/119=6.7%	111/119=93.3%
Total	9/302=3%	293/302=97%

Fonte: PORTVIX, 2002.

³² Informação repassada pela Editora Abril.

Como já foi observado no início desta seção, este grupo não foi codificado na amostra da escrita, visto que todos os autores dos textos incluídos na base de dados se inserem no mesmo fator, que é o de nível superior de ensino.

Este fator, *a priori*, como se viu no capítulo 1, na seção 1.2, foi o único de caráter não linguístico apontado hipoteticamente como possível influenciador na variação sintática do fenômeno deste estudo. Nesse sentido, esperava-se que, sob a influência desse grupo de fatores, houvesse maior número de ocorrências de finais desenvolvidas à medida que o nível de escolarização fosse aumentando. Essa expectativa decorre do fato de que a língua portuguesa trabalhada na escola está alinhada ao conhecimento das gramáticas normativas, que, com seus modelos literários, tendem a enfatizar os tipos de orações subordinadas iniciadas por conjunções e locuções conjuntivas – no caso em análise, a única encontrada foi *para que*. Isso significa dizer que, se a ênfase é nesse tipo de emprego, ela recai sobre a variante final desenvolvida, o que poderia levar o mais escolarizado a usá-las mais que o menos escolarizado.

Mesmo com um baixo número de ocorrências, apenas 9, a expectativa a respeito da relação entre a frequência da variante desenvolvida e o aumento da escolarização dos informantes foi confirmada nas 19 células investigadas na base de dados do PORTVIX. É passível de observação que há um movimento crescente na frequência de finais desenvolvidas à medida que aumenta a escolarização – fundamental: 0%; médio: 1.1%; superior; 6.7%.

Da mesma forma que ocorreu no grupo de fatores *faixa etária*, também houve efeito categórico neste, pois, no nível fundamental de escolarização, todas as ocorrências foram de finais reduzidas. Para que se pudesse fazer a rodada de pesos relativos, os dados desse fator foram amalgamados com os referentes ao ensino médio, processo que será detalhado na próxima seção.

No tocante à amostra da escrita, pelo fato de os dados estarem todos vinculados ao fator *nível superior de ensino*, a expectativa era a de encontrar um número de finais desenvolvidas superior ao da amostra da fala, o que na prática aconteceu, já que 16% dos dados das reportagens correspondem à variante marcada.

No entanto, para que essa questão não ficasse restrita a uma interpretação sem a devida apuração estatística, que perpassa toda a metodologia desta pesquisa, foi empreendida uma codificação geral dos dados encontrados nos dois *corpora*, de forma que, mantida a variável

dependente, incorporou-se outra variável independente às 8 já existentes: a *modalidade de expressão da língua*, cujos fatores são *fala* e *escrita*. Os resultados obtidos nessa fusão podem ser vistos na Tabela 12:

TABELA 12

Frequências absoluta e relativa de todos os dados, com base no grupo de fatores
modalidade de expressão da língua

Modalidade	Desenvolvidas	Reduzidas
<i>(a) Fala</i>	9/302=3%	293/302=97%
<i>(b) Escrita</i>	45/282=16%	237/282=84%
Total	54/584=9.2%	530/584=90.8%

Fonte: PORTVIX, 2002; Superinteressante, 2003, 2004.

Por meio desses resultados, comprova-se que a língua falada, no gênero entrevista, é favorecedora da ocorrência das finais reduzidas, pois aumenta a sua média de 90.8% para 97%, em meio aos 584 dados coletados e analisados neste estudo. Por outro lado, as desenvolvidas tendem a emergir discursivamente quando o usuário faz uso da modalidade escrita da língua, no gênero reportagem. Nesse caso, a média global de finais desenvolvidas, que é de 9.2%, sobe para 16%, o que revela um aumento de quase 7 pontos percentuais. Para corroborar essa forte tendência, é importante informar que na aferição dos pesos relativos com a presença deste grupo de fatores, obteve-se 0.78 de efeito da escrita e 0.24 da fala para as finais desenvolvidas.

A partir deste ponto, serão relatados os resultados das rodadas dos pesos relativos obtidos por meio do programa computacional Varbrul (SANKOFF; SMITH; TAGLIAMONTE, 2005), assim como os grupos de fatores que o mesmo programa selecionou por terem grau de relevância estatística para a compreensão da variação do fenômeno linguístico em estudo.

5.3 Grupos de fatores relevantes estatisticamente

Conforme a descrição feita anteriormente, alguns fatores dos grupos definidos para esta análise apresentaram efeito categórico – 100% dos casos recaindo em apenas uma das variantes da variável dependente. Particularmente, esse efeito categórico ocorreu sempre da seguinte maneira: nenhuma ocorrência vinculada às finais desenvolvidas e todas às finais reduzidas. Em nenhum grupo de fatores aconteceu a relação inversa dessa distribuição.

No entanto, para seguir com a análise pelo programa Varbrul para além da demonstração de frequências absolutas e relativas, e para que sejam identificados aspectos mais confiáveis de análise do ponto de vista estatístico, o programa desencadeia a chamada *rodada de pesos relativos*³³. Essa rodada (cálculo) indica ao pesquisador as variáveis independentes ou grupos de fatores que tendem a influenciar mais fortemente a alternância entre as variantes da variável dependente definida como objeto de investigação. Nesse sentido, para que os pesos relativos sejam apurados e os grupos de fatores indicados, não pode haver grupos com fatores que apresentem efeito categórico, como se viu nas seções anteriores. Em função disso, alguns ajustes na identificação dos dados precisaram ser feitos.

Tais ajustes só foram necessários em dois grupos de fatores, de natureza extralinguística, aplicados à amostra da língua falada. Isso significa que nos dados levantados no *corpus* da escrita não ocorreu quantificação com efeito categórico.

O primeiro dos grupos ajustados foi o referente à *faixa etária* dos informantes do PORTVIX. Na Tabela 10, da seção 5.2.2, pôde-se observar que, quando se tratou dos fatores que compreendiam as faixas etárias de *7 a 14 anos* e *50 anos ou mais*, só houve dados codificados numa das variantes, o que significa dizer que, dos 302 dados levantados na amostra da fala, 11.6% (35 casos) foram identificados em falantes de 7 a 14 anos, e 24.5% (74 casos) em falantes de 50 anos em diante, sendo todos exemplos de finais reduzidas. Em virtude dessas ocorrências, decidiu-se agrupar, por meio de um mecanismo chamado *amalgamação de fatores* (GUY; ZILLES, 2007, p. 188), os dados referentes à faixa de *7 a 14* com os de *15 a 25 anos*, e os de *50 anos em diante* com os de *26 a 49 anos*. O resultado desses agrupamentos foi que, no lugar de 4 fatores, esta variável independente passou a ter apenas 2. Essas amalgamações foram feitas com base na proximidade etária de um grupo em relação ao outro.

³³ As rodadas de peso relativo estão dispostas nos anexos deste texto.

Entretanto, amalgamar os fatores do grupo faixa etária poderia trazer um resultado pouco preciso e que talvez não condissesse com os reais efeitos dessa variável sobre o fenômeno linguístico em foco. Em função disso, tomou-se a decisão de se fazer uma nova rodada de pesos relativos sem os dados codificados para os dois fatores nos quais não houve distribuição binária. Desse modo, com a exclusão dos 35 dados da faixa que vai dos 7 aos 14 anos e dos 74 referentes à faixa dos 50 anos em diante, a nova rodada contou com apenas 193 dados. Ainda assim, os resultados foram praticamente os mesmos obtidos com a presença de todos os 302 dados inicialmente codificados, pois as variáveis independentes selecionadas por sua relevância estatística foram mantidas, e permaneceram as mesmas tendências probabilísticas indicadas pelos pesos relativos aferidos. Em virtude dessa constatação, continuaram a ser utilizados para esta análise os resultados que contêm todos os dados codificados. Esta rodada pode ser conferida no Anexo 3, ao final do trabalho.

O outro grupo de fatores no qual se identificou efeito categórico foi o de *nível de escolarização*. A categoricidade se deu no fator *ensino fundamental*, pois todas as ocorrências referentes a esse nível eram casos de finais reduzidas, o que corresponde a 90 dos 302 dados, de acordo com o que se viu na Tabela 11, da seção 5.2.3. Por isso, aplicou-se o mesmo procedimento feito no grupo *faixa etária*, isto é, foram amalgamados os fatores *ensino fundamental* e *médio*, o que se justifica pela proximidade entre esses níveis de escolarização e também pelos efeitos que ambos apresentaram. Assim, passaram a constituir esta variável apenas dois fatores e não mais três: *ensino fundamental* e *médio*, amalgamados, e *ensino superior*.

Somente depois desse procedimento é que os pesos relativos puderam ser calculados. Guy e Zilles (2007, p. 239) definem peso relativo como

um valor calculado pelo Varbrul (com base em um conjunto de dados) que indica o efeito deste fator sobre o uso da variante investigada neste conjunto. O valor dos pesos recai sempre no intervalo entre zero e um (0-1), em que um valor zero indica que tal variante nunca acontece quando este fator está presente, e um valor de 1 indica que tal variante sempre ocorre quando o fator está presente.

Assim, o Varbrul pôde selecionar, por meio da análise de relevância e significância estatística, alguns grupos de fatores que probabilisticamente influenciam no uso das variantes analisadas nesta pesquisa.

Em relação aos dados da língua falada, o programa selecionou como relevantes estatisticamente os seguintes grupos: o (2), *características sintático-semânticas do sujeito da oração principal*, o (4), *explicitação do sujeito na adverbial final*, ambos de natureza linguística, e o (8), *nível de escolarização*, de natureza extralinguística. Sobre os dados da língua escrita, o Varbrul selecionou apenas grupos de caráter linguístico: o (1), *correferencialidade do sujeito da adverbial final*, e o (4), *explicitação do sujeito na adverbial final*.

Nas Tabelas 13 e 14, a seguir, são retomadas as frequências referentes às variáveis selecionadas, bem como os pesos relativos de cada um de seus fatores. Observa-se que são apresentados os números relacionados à variante *final desenvolvida*, de maneira que, para interpretar os valores referentes à variante *final reduzida*, basta fazer o cálculo complementar, tanto para os números absolutos e porcentagens, quanto para os pesos relativos.

TABELA 13

Atuação das variáveis selecionadas por relevância estatística no uso das finais desenvolvidas nos dados da fala

Variáveis selecionadas	Frequências	Peso relativo
(2) Características sintático-semânticas do sujeito da oração principal		
(a) <i>sujeito controlador/animado</i>	3/276=1.1%	0.44
(b) <i>sujeito não controlador/inanimado/oracional</i>	6/26=23.1%	0.94
(4) Explicitação do sujeito na adverbial final		
(a) <i>sujeito explícito</i>	8/74=10.8%	0.91
(b) <i>sujeito não explícito</i>	1/228=0.4%	0.32
(8) Nível de escolarização		
(a) <i>ensino fundamental e médio</i>	1/183=0.5%	0.25
(b) <i>ensino superior</i>	8/119=6.7%	0.84
Total	9/302=3%	

Fonte: PORTVIX, 2002.

Nesta tabela, estão expostas as variáveis independentes que apresentaram, após as rodadas de pesos relativos, relevância estatística para a compreensão do fenômeno variável deste trabalho nos dados da fala. Essa visão generalizada facilita perceber a atuação de cada fator em termos do favorecimento às variantes, bem como auxilia na configuração da área de variação do fenômeno linguístico aqui observado/analísado. De forma sintética, o que se tem é o seguinte panorama: as finais desenvolvidas são fortemente favorecidas quando atuam os fatores *sujeito não controlador*, *sujeito explícito* e *nível superior de escolarização*, com 0.94, 0.91 e 0.84 de peso relativo, respectivamente. O “termômetro” dos pesos relativos deve ser compreendido tomando-se como referência o valor que está acima ou abaixo de 0.50. Como os pesos relativos dos fatores selecionados estão acima desse valor e mais próximos de 1(um), aumentam as chances de ocorrência da variante final desenvolvida quando houver a confluência desses três fatores. Assim, embora a margem de variação na fala seja pequena, o efeito das variáveis estatisticamente significativas é robusto.

Uma vez que todos os grupos de fatores selecionados são binários, é possível depreender, por meio da exposição dos valores dos pesos relativos, que os fatores não especificados – *sujeito controlador*, *sujeito não explícito* e *níveis fundamental e médio de escolarização* – favorecem a ocorrência das finais reduzidas, já que as diferenças entre os dois pesos são todas muito altas. É interessante observar, nesse sentido, que os pesos relativos confirmaram a tendência que havia sido verificada quando se analisou, nas seções anteriores, o aumento da frequência relativa (%) em relação à média global das ocorrências.

Como é verificável na Tabela 13, a única variável independente de natureza social selecionada com significância do ponto de vista estatístico foi o *nível de escolarização*, para os dados que compõem a amostra da língua falada. No capítulo 1, na seção 1.2, ao se levantarem as hipóteses para esta pesquisa, falou-se a respeito da possível influência desse fator, sobretudo como favorecedor de um número um pouco maior de finais desenvolvidas à medida que o nível de escolarização dos informantes fosse aumentando.

Conforme já foi comentado na seção 5.2.3, houve amalgamação dos fatores *ensino fundamental e médio*, pois, como foi demonstrado na Tabela 11, havia ocorrido efeito categórico para falantes com ensino fundamental, com 100% de finais reduzidas, e frequência de 1.1% para os com ensino médio – assim, a fala que exibe mais variação é a dos usuários com ensino universitário. No que tange aos números, as finais desenvolvidas são favorecidas pelo fator *ensino superior*, com 0.84 de peso relativo, e as reduzidas pelos outros dois níveis

de ensino, o *fundamental* e o *médio*, com 0.75 de peso relativo, um índice bastante significativo do ponto de vista estatístico, observando-se este segundo peso de forma complementar ao peso relativo de 0.25, que é desfavorecedor das finais desenvolvidas para o ensino fundamental e médio.

A hipótese inicial a respeito da influência da escolarização, portanto, se confirma com esses resultados. É, então, um fator que favorece o uso da variante mais marcada na escala do processo de redução das orações finais.

Deste ponto em diante, verifica-se a seleção para os dados da língua escrita:

TABELA 14

Atuação das variáveis selecionadas por relevância estatística no uso das finais desenvolvidas nos dados da escrita

Variáveis selecionadas	Frequências	Peso relativo
(1) Correferencialidade do sujeito da adverbial final		
<i>(a) sujeito correferente</i>	4/191=2.1%	0.36
<i>(b) sujeito parcialmente correferente</i>	8/27=29.6%	0.81
<i>(c) sujeito não correferente</i>	33/64=51.6%	0.76
(4) Explicação do sujeito na adverbial final		
<i>(a) sujeito explícito</i>	36/42=85.7%	0.97
<i>(b) sujeito não explícito</i>	9/240=3.8%	0.35
Total	45/282=16%	

Fonte: Superinteressante, 2003, 2004.

Nesta tabela, estão as duas únicas variáveis independentes selecionadas pelo Varbrul, para os dados extraídos das reportagens, como importantes do ponto de vista estatístico para a análise do fenômeno variável desta pesquisa. São dois grupos de fatores altamente correlacionáveis porque tratam do mesmo termo da oração, que é o sujeito.

Pelos números, a retomada de algum termo da oração à qual a final está vinculada, codificada no fator *sujeitos parcialmente correferentes*, é o que mais favorece a ocorrência de finais

desenvolvidas, com peso relativo de 0.81. Não muito atrás, aparece o fator *sujeitos não correferentes*, com 0.76 de peso relativo, também favorecendo fortemente essa variante. Na verdade, ao se fazer uma comparação estrutural entre os três fatores que compõem o grupo, será atestado que o *status* sintático-semântico do que aqui se está chamando de correferência parcial e não correferência é muito semelhante, daí os pesos relativos para os dois fatores estarem tão aproximados e indicarem a mesma tendência. É uma interessante prova de que o estatístico está refletindo o linguístico.

Na mesma linha, por conseguinte, vem a *explicitação do sujeito* apontando, assim como ocorreu nos dados da fala, o favorecimento da ocorrência de finais desenvolvidas, com 0.97 de peso relativo, o que serve para confirmar a tendência já anteriormente verificada na apresentação da frequência relativa, que é de 85.7% de desenvolvidas quando o sujeito se explicita, algo que aumenta sobremaneira a sua média, que é de 16% do total de dados coletados.

Seguindo o mesmo raciocínio explanado na atuação das variáveis da fala, deve-se entender que as finais reduzidas são favorecidas quando estão presentes os seguintes fatores: *sujeitos correferentes* e *não explícitos*, ambos com peso relativo de 0.65.

5.4 Considerações sobre a correferencialidade e explicitação do sujeito da adverbial final

Das cinco variáveis independentes de natureza linguística eleitas para o estudo do fenômeno variável deste trabalho, duas delas trazem características comuns entre si, uma vez que tratam de um termo de mesma função sintática, que é o sujeito. De acordo com a ordem com que foram anteriormente apresentadas, essas variáveis são: (1) *correferencialidade* e (4) *explicitação do sujeito da adverbial final*. Além disso, esta seção dedicada a elas se deve ao fato de a variável independente (4) ter sido a única apontada com significância estatística para a análise nas duas amostras. Nada mais sensato, então, que fazer uma correlação entre duas variáveis tão imbricadas linguisticamente.

Para iniciar essa discussão, tome-se a Tabela 15, na qual está especificado o cruzamento dos resultados obtidos nos *corpora* da fala e da escrita para a correferencialidade e a explicitação do sujeito.

TABELA 15

Cruzamento entre (1) correferencialidade e (4) explicitação do sujeito da oração final

Correferencia- lidade	Explicitação	Fala		Escrita	
		Desenvolvidas	Reduzidas	Desenvolvidas	Reduzidas
<i>subj. correferente</i>	<i>explícito</i>	0/28=0%	28/28=100%	1/1=100%	0/1=0%
	<i>não explícito</i>	1/184=1%	183/184=99%	3/190=2%	187/190=98%
<i>subj. parc. correferente</i>	<i>explícito</i>	1/6=17%	5/6=83%	6/6=100%	0/6=0%
	<i>não explícito</i>	0/30=0%	30/30=100%	2/21=10%	19/21=90%
<i>subj. não correferente</i>	<i>explícito</i>	7/40=18%	33/40=82%	29/35=83%	6/35=17%
	<i>não explícito</i>	0/14=0%	14/14=100%	4/29=14%	25/29=86%
Total		9/302=3%	293/302=97%	45/282=16%	237/282=84%

Fonte: PORTVIX, 2002; Superinteressante, 2003, 2004.

Diante desta tabela, alguns aspectos no que tange à distribuição e à frequência dos dados nos dois *corpora* podem ser observados e merecem seus devidos comentários. O primeiro destaque são alguns contextos de invariância identificados no cruzamento das duas variáveis independentes. Na fala, os casos de categoricidade (100% das ocorrências) referem-se a dados de finais reduzidas e que estão na convergência dos seguintes fatores:

- a) *sujeito correferente + sujeito explícito*, com 28 ocorrências, como em:

[...] você tem que passar isso pra criança e começa de pequeno... PRA VOCÊ NÃO VER O QUE TÁ ACONTECENDO AÍ HOJE [...]

- b) *sujeito parcialmente correferente + sujeito não explícito*, com 30 ocorrências, como em:

[...] você já acorda agora tem gente assim como o moço ali embaixo ele não:: gosta de que acorda ele de madrugada PRA VENDER CAMARÃO aí ele dá esporro na pessoa [...]

- c) *sujeito não correferente + sujeito não explícito*, com 14 ocorrências, como em:

[...] se PRA FEDERALIZAR MESMO... aquele negócio da Farmácia tá difícil, você imagine... hoje... com esse monte de curso... esse monte de coisa [...]

Sendo assim, em se tratando dos dados retirados das 19 células da amostra PORTVIX, os resultados registram que não houve variação entre finais desenvolvidas e reduzidas quando essas três combinações estavam presentes. Em outros termos, essas combinações não possibilitaram a emersão discursiva das finais desenvolvidas.

Viu-se, na seção 5.3, que o grupo da correferencialidade não foi selecionado como estatisticamente relevante para a variação na fala. Nos níveis 1 e 2 da rodada de pesos relativos, esta variável estava dentro dos padrões de significância, com potencial para ser selecionada. No entanto, no nível 3, quando o grupo (4) – o da explicitação – é agregado à análise conjunta com o grupo (1) – o da correferencialidade – e o grupo (2) – o dos traços sintático-semânticos do sujeito da oração principal –, o nível de significância dessa convergência fica bem acima do aceitável para a seleção estatística. Isso pode ser explicado justamente pelo fato de haver esses três contextos de invariância quando estão conjugadas a (não) correferencialidade e a (não) explicitação.

No que diz respeito a esse mesmo aspecto – a invariância na área de confluência das duas variáveis independentes em questão –, a amostra da escrita revelou divergências em relação à da fala, visto que, inversamente a este *corpus*, aquele revelou efeito categórico com dados somente de orações finais desenvolvidas em duas áreas do cruzamento. São elas:

- a) *sujeito correferente + sujeito explícito*, com 1 ocorrência apenas, que é:

Foi aí que se instituiu a monogamia. PARA QUE O HOMEM PUDESSE TER CERTEZA DE QUE O FILHO GERADO ERA SEU, teria que garantir que a mulher não tinha feito sexo com outros.

- b) *sujeito parcialmente correferente + sujeito explícito*, com 6 ocorrências, como em:

[...] o muco retém as partículas odoríferas PARA QUE ELAS POSSAM SER ANALISADAS PELOS NEURÔNIOS RECEPTORES.

Vistas as zonas em que houve efeito categórico, é importante discorrer um pouco sobre aquelas que indicam uma distribuição mais equilibrada dos dados – lembrando que essa observação está sendo feita com base na zona de cruzamento das duas variáveis independentes linguísticas que, segundo os resultados, tendem a influenciar mais fortemente a variação entre as finais desenvolvidas e reduzidas. Nesse sentido, na Tabela 15, nas linhas em que estão associados os fatores *sujeito correferente + sujeito não explícito* e *sujeito não*

correferente + sujeito explícito, observa-se que houve ocorrência de dados nas duas variantes da variável dependente, na fala e na escrita. Essa constatação é interessante na medida em que essas combinações são aquelas das quais se espera o maior efeito em termos de motivação, tanto econômica, quanto icônica (HAIMAN, 1983): em ambas as associações, o primeiro fator tende a motivar, por economia, o segundo; na primeira associação, uma menor codificação tende a refletir uma menor estrutura oracional, ou seja, uma oração reduzida, ao passo que, na segunda, uma maior codificação tende a refletir uma maior estrutura oracional, ou seja, uma oração desenvolvida, o que é resultado de uma relação iconicamente motivada.

Outro aspecto interessante que a Tabela 15 expõe são as tendências convergentes e divergentes no que tange às confluências da (não) correferencialidade e da (não) explicitação do sujeito da adverbial final. Observando-se o cruzamento de qualquer fator do grupo da correferencialidade com o fator *sujeito não explícito*, chega-se à conclusão de que, nas duas bases de dados, essa associação apresenta resultados muito próximos quando se comparam as frequências em termos de favorecimento e desfavorecimento às orações desenvolvidas ou reduzidas. Por exemplo, quando estão associados os fatores *sujeito correferente* com *sujeito não explícito*, confirma-se a tendência à redução oracional – em função do princípio da economia linguística, ou seja, não codificar aquilo que já é discursivamente previsível (HAIMAN, 1983). Na fala, o sujeito da oração final não se explicita nas 183 das 211 vezes em que a oração reduzida apresenta este mesmo sujeito correferente ao da oração principal ou ao de qualquer oração à qual a final estiver vinculada, o que implica um percentual de 86.7%. Há, ainda na fala, 1 (um) único caso de final desenvolvida com a convergência entre não explicitação e correferencialidade, cuja transcrição é a seguinte:

[...] *o ensino público, ele* deve ser melhorado desde a base PRA QUE SEJA TÃO BOM QUANTO AS ESCOLAS PARTICULARES [...]

Por sinal, esta ocorrência é a única de final desenvolvida com sujeito não explícito entre todos os dados coletados da amostra PORTVIX.

Na escrita, a tendência é a mesma, pois das 187 orações finais reduzidas com sujeito correferente, nenhum deles se explicitou. Contudo, os números revelam que, dos 4 casos de finais desenvolvidas com sujeito correferente, em 3 deles não há explicitação – o único caso em que se expressa o sujeito foi exposto anteriormente, quando se falou das áreas de

invariância. Assim sendo, examine-se novamente esse dado, e na sequência os outros 3, para que sejam conhecidos seus aspectos estruturais e discursivos:

Foi aí que se instituiu a monogamia. PARA QUE O HOMEM PUDESSE TER CERTEZA de que o filho gerado era seu, teria que garantir que a mulher não tinha feito sexo com outros.

Ao longo deste trabalho, por várias vezes se disse que a expectativa é de que a final desenvolvida venha com sujeito expreso. Mas o que não se espera é que ele seja correferente ao da principal. Neste período misto, em que se narram jornalisticamente os primórdios das relações monogâmicas nas famílias, a oração final ocorre anteposta à principal – que é a posição marcada das finais quando se olham os percentuais gerais (73% de posposição e 27% de anteposição na amostra da escrita). Somando-se a isso o fato de se tratar de escrita monitorada e revisada, com vistas a eliminar certas marcas de redundância nas formas linguísticas, o sujeito expreso na oração anteposta leva à sua não explicitação na oração principal, pelo fato de haver correferência desse sujeito – novamente a atuação do princípio da economia linguística. Assim, a explicitação do sujeito neste exemplo de final desenvolvida parece estar relacionada à anteposição da oração final. E o fato de haver a correferência do sujeito não inibiu a sua expressão, também em função da ordem marcada. Entretanto, é importante frisar que o fator *sujeito correferente* desfavorece a emersão discursiva das finais desenvolvidas na escrita (0.36 de peso relativo, cf. Tabela 14).

Além disso, numa mudança potencial da ordem, o sujeito “homem” teria de vir expreso na oração principal, para que fosse garantida a coesão textual, o que levaria ao seu cancelamento na final: *O homem teria que garantir que a mulher não tinha feito sexo com outros, PARA QUE PUDESSE TER CERTEZA de que o filho gerado era seu.* Ainda assim, vê-se que é possível a ocorrência da final desenvolvida, mas as condições para a sua realização sob a forma reduzida são favoráveis: *O homem teria que garantir que a mulher não tinha feito sexo com outros, PARA PODER TER CERTEZA de que o filho gerado era seu.*

Os outros três casos de finais desenvolvidas são de correferencialidade do sujeito, sem haver a sua explicitação na oração final:

(1) Galinhas poedeiras vivem espremidas sob luz quase ininterrupta PARA QUE COMAM E BOTEM OVOS SEM PARAR.

(2) Ponha-se no lugar de um frango que vive amontoado sob luz artificial quase ininterrupta, equilibrando-se no arame do fundo da gaiola e que tem seu bico cortado com uma lâmina quente, para evitar que ele selecione a comida que lhe é dada e PARA QUE, DE NERVOSO, NÃO ATAQUE E MATE OS OUTROS FRANGOS, até que, um dia, uma descarga elétrica o ponha para dormir enquanto uma lâmina corta seu pescoço.

(3) Dawkins propôs a palavra "meme" para designar essa nova entidade. O termo vem do grego mimeme [imitação], reduzido a duas sílabas PARA QUE SOASSE PARECIDO COM "GENE".

Nestas três ocorrências, encontram-se orações finais desenvolvidas que não atendem aos modelos que mais se esperam delas, que estejam próximas de seu protótipo (GIVÓN, 1995). O elemento mais importante que contribui para isso é a não explicitação do sujeito, motivada por sua correferencialidade com o sujeito da oração principal. No entanto, na ordem em que estão dispostos os exemplos, os sujeitos “galinhas”, “frango” e “termo” apresentam o traço semântico não controlador sobre a finalidade expressa nas orações desenvolvidas. Essa propriedade semântica do sujeito foi monitorada no grupo de fatores (2), quando, em função dos resultados encontrados, foi verificado que sua ocorrência é favorecedora da variante final desenvolvida, apesar de a variável não ter sido selecionada com significância estatística – no caso da escrita, a média de desenvolvidas sobe de 16% para 26.2% na presença do fator sujeito não controlador.

As últimas considerações a respeito da correlação entre (não) correferencialidade e (não) explicitação do sujeito na adverbial final serão feitas a partir da comparação dos resultados da base de dados da fala deste trabalho (PORTVIX) com os obtidos na pesquisa de Finck (2000), com dados do VARSUL, apenas em caráter de distribuição dos dados. Observe-se a Tabela 16, em que é recuperada parte dos resultados da Tabela 15 com o fito de melhor visualizar a comparação proposta.

TABELA 16

Cruzamento entre (1) correferencialidade e (4) explicitação do sujeito nas finais reduzidas nos dados da fala: Vitória (ES) versus Florianópolis (SC)

Correferencialidade	Vitória (PORTVIX)		Florianópolis (VAR SUL)	
	<i>Suj. explícito</i>	<i>Suj. não explícito</i>	<i>Suj. explícito</i>	<i>Suj. não explícito</i>
<i>correferente</i>	28/211=13.3%	183/211=86.7%	12/118=10%	106/118=90%
<i>parcial. correferente</i>	5/35=14.3%	30/35=85.7%	3/17=18%	14/17=82%
<i>não correferente</i>	33/47=70.2%	14/47=29.8%	30/47=64%	17/47=36%
Total	66/293=23%	227/293=77%	45/182=25%	137/182=75%

Fonte: PORTVIX, 2002; FINCK, 2000.

Embora o total de dados coletados nas duas amostras seja diferente, é verificável na Tabela 16 que as tendências nas duas capitais são as mesmas. A combinação *sujeito correferente + sujeito não explícito* – que é a associação prototípica nas finais reduzidas – foi a que apresentou maior frequência relativa: 86.7% em Vitória e 90% no em Florianópolis. Também pode ser observado que a média de não explicitação, que é de 77% em Vitória e de 75% em Florianópolis, aumenta significativamente, nas duas capitais, na presença do fator *sujeito correferente*.

Ao se esquematizar um *continuum* que vai da maior para a menor frequência a partir desse cruzamento com os dados das finais reduzidas, obtém-se a seguinte configuração para Vitória e Florianópolis, respectivamente:

1. suj. correferente + não explícito (86.7% e 90%);
2. suj. parc. correferente + não explícito (85.7% e 82%);
3. suj. não correferente + explícito (70.2% e 64%);
4. suj. não correferente + não explícito (29.8% e 36%);
5. suj. parc. correferente + explícito (14.3% e 18%);
6. suj. correferente + explícito (13.3% e 10%).

Essa configuração é notoriamente harmoniosa, no sentido de que a escala começa com uma associação altamente motivada – mais correferência, menos explicitação – e termina com uma que tem correlação fraca, corroborada pelas frequências mais baixas encontradas no cruzamento feito.

Essa analogia vem confirmar o que já se afirmou na seção referente às variáveis sociais sobre o fenômeno linguístico investigado nesta dissertação: ele não é fortemente caracterizado por influências externas ao sistema linguístico, já que apenas a variável referente à escolarização apresentou relevância estatística na amostra da fala. Como essa comparação considera uma variável geográfica, confirma-se não ser de caráter regional o fenômeno da variação sintática das adverbiais finais, mas, sim, de abrangência ampla na língua portuguesa, nesse caso, a língua portuguesa falada, mais precisamente no gênero entrevista nos moldes labovianos.

Ainda sobre a configuração sintática do sujeito das orações finais, outras pesquisas na área da linguística, tanto no âmbito da coordenação, quanto da subordinação, ao investigarem os diferentes graus de integração existentes entre as orações, procuraram analisar e entender a influência que a correferência e a explicitação do sujeito têm sobre esse mecanismo sintático. Nesse sentido, Görski (2001) trabalha com diferentes tipos de orações adverbiais, entre elas as finais. Partindo de propostas da teoria funcionalista sobre os graus de integração das orações, como as de Haiman e Thompson (1984), Lehmann (1988) e Givón (1990), Görski (2001) destaca, como um dos aspectos do seu trabalho, as possíveis correlações entre a não explicitação do sujeito, a correferencialidade e a redução de orações – o que nesta dissertação é assumido como uma motivação dupla (HAIMAN, 1983): *econômica*, quando não se explicita o sujeito por ser ele correferente, e *icônica*, pelo fato de a menor codificação linguística resultante da economia linguística fazer emergir discursivamente uma estrutura também menor, que é a oração reduzida.

Em seu trabalho, Görski retoma dados de outra pesquisa sua – Görski (2000) –, adicionando-os a outros de diferentes bases de pesquisa, inclusive os de Fink (2000). Nesses dados, estão englobados vários tipos de orações com traço semântico de finalidade – não apenas as adverbiais finais – e que tenham a estrutura “para + infinitivo”. Ainda assim, os resultados obtidos nesta dissertação estão alinhados com os encontrados por Görski: “[...] em 79% dos dados analisados não há sujeito explícito. Desses, 60% apresentam identidade de sujeito [...]. Portanto, sujeito não explicitado e correferencial são fatores preponderantes nas construções

para INF” (GÖRSKI, 2001, p. 169). Ou seja, são preponderantes para que as orações sejam reduzidas.

Na próxima seção, será proposta uma escala para as orações finais, com base no princípio da marcação linguística.

5.5 Uma escala para as orações adverbiais finais

Após entender melhor o comportamento sintático-semântico das orações finais, em termos de seu uso variável nas diferentes amostras analisadas, propõe-se nesta seção uma sequência que mostra essas orações numa escala baseada na configuração linguística que apresentam, da mais para a menos frequente.

Essa amostragem escalar leva em consideração o princípio da marcação linguística. De acordo com Givón (1995), a identificação de uma forma linguisticamente marcada ou não deve levar em conta aspectos como a complexidade formal e cognitiva da estrutura linguística, bem como a frequência com que ela ocorre na língua em uso. Assumiu-se neste trabalho – conforme o capítulo 3 – a noção de marcação não apenas quando uma forma não marcada se opõe a outra marcada, mas sim numa perspectiva gradiente de casos menos marcados a mais marcados.

Como já se confirmou, de forma ampla, as finais desenvolvidas são o caso marcado, e as reduzidas, o não marcado. Sendo assim, o objetivo aqui é ir para além dessa oposição dicotômica, de forma a apresentar uma escala de acordo com a combinação de determinados aspectos linguísticos. Nesse sentido, será tomado como principal critério para a atribuição dessa escalaridade a frequência identificada na apuração dos resultados, tendo em vista os grupos de fatores de natureza linguística que se mostraram mais imbricados e dependentes entre si, quantitativa e qualitativamente. Os grupos a que se está fazendo referência são: (1) correferencialidade do sujeito da adverbial final e (4) explicitação do sujeito da adverbial final.

Antes da apresentação do quadro com os níveis da escala, é importante a visualização da Tabela 17, na qual está a distribuição dos dados em termos de sua frequência e que serviu de referência para a organização da escala para as orações finais.

TABELA 17

Cruzamento entre (1) correferencialidade e (4) explicitação do sujeito da oração final

(1)	(4)	Fala		Escrita	
		Desenvolvidas	Reduzidas	Desenvolvidas	Reduzidas
<i>subj. correferente</i>	<i>Explícito</i>	0/1=0%	28/211=13.3%	1/4=2.5%	0/187=0%
	<i>não explícito</i>	1/1=100%	183/211=86.7%	3/4=97.5%	187/187=100%
<i>subj. parc. correferente</i>	<i>Explícito</i>	1/1=100%	5/35=14.3%	6/8=75%	0/19=0%
	<i>não explícito</i>	0/1=0%	30/35=85.7%	2/8=25%	19/19=100%
<i>subj. não correferente</i>	<i>Explícito</i>	7/7=100%	33/47=70.2%	29/33=87.9%	6/31=19.4%
	<i>não explícito</i>	0/7=0%	14/47=29.8%	4/33=12.1%	25/31=80.6%

Fonte: PORTVIX, 2002; Superinteressante, 2003, 2004.

Com base nesta distribuição, é proposto o quadro a seguir, em que é apresentada uma escala comparativa entre os dados da fala e da escrita, com base na associação entre (1) e (4) para as orações finais. Deve ser observado que não se trata uma escala que opõe desenvolvidas a reduzidas, mas que engloba as ocorrências das orações finais como um todo.

É importante anotar que a associação entre as variáveis indicadas foi feita por meio do mecanismo de Tabulação Cruzada (*Cross Tabulation*), propiciado pelo programa computacional Varbrul. Sendo assim, em caráter de frequência decrescente, entre finais desenvolvidas e reduzidas, a configuração é esta:

QUADRO 3**Aplicação do princípio da marcação às orações finais, com base na associação entre correferencialidade e explicitação do sujeito da adverbial final, na fala e na escrita**

FALA	ESCRITA
<p>1. Reduzidas com sujeito correferente não explícito, com 183 casos.</p> <p>Ex.: [...] ainda mais que eu moro longe, entendeu, pegar dois ônibus PRA IR PRA LÁ EH MEIO DIFÍCIL [...]</p>	<p>1. Reduzidas com sujeito correferente não explícito, com 187 casos.</p> <p>Ex: Ao chegar à escola, ele passa por uma seleção PARA SABER EM QUE CASA INGRESSARÁ [...]</p>
<p>2. Reduzidas com sujeito não correferente explícito, com 33 casos.</p> <p>Ex.: [...] eu creio que deveria ser frequentemente, né? pro pessoal tá sempre em alerta, correto, PRA ELES FICAREM CIENTES DO QUE ELES ESTÃO FAZENDO.</p>	<p>2. Desenvolvidas com sujeito não correferente explícito, com 29 casos.</p> <p>Ex.: PARA QUE OS ETs CONHEÇAM AS MENSAGENS, as sondas devem ser capturadas no espaço [...]</p>
<p>3. Reduzidas com sujeito parc. correferente não explícito, com 30 casos. (não houve dados de desenvolvidas).</p> <p>Ex.: [...] e o filho dele... era quem levou ele PRA VER A MÃE [...]</p>	<p>3. Reduzidas com sujeito não correferente não explícito, com 25 casos.</p> <p>Ex.: PARA TESTAR A TOXICIDADE DE UM PRODUTO DE LIMPEZA, COSMÉTICOS E MEDICAMENTOS, 6 a 9 animais são forçados a ingerir doses crescentes da substância [...].</p>
<p>4. Reduzidas com sujeito correferente explícito, com 28 casos (não houve dados de desenvolvidas).</p> <p>Ex.: [...] você tem que passar isso pra criança e começa de pequeno... PRA VOCÊ NÃO VER O QUE TÁ ACONTECENDO AÍ HOJE...</p>	<p>4. Reduzidas com sujeito parc. correferente não explícito, com 19 casos.</p> <p>Ex.: Ela para de agir assim que o intestino está liberado, abandonando você à própria sorte PARA RESISTIR ÀS TENTAÇÕES.</p>
<p>5. Reduzidas com sujeito não correferente não explícito, com 14 casos (não houve dados de desenvolvidas).</p> <p>Ex.: [...] eu 'tô pagando esse atendimento e tal ... e ela ela vai me dar alguns textos PRA LER E ESCREVER né?</p>	<p>5. Reduzidas com sujeito não correferente explícito, com 6 casos.</p> <p>Ex.: A vitela é a carne de um bezerro anêmico que passa seus cinco meses de vida em um cercado minúsculo, impedido de se mover, PARA A CARNE FICAR MACIA.</p>

<p>6. Desenvolvidas com sujeito não correferente explícito, com 7 casos.</p> <p>Ex.: [...] eu tenho que gerenciar e organizar ... fazer um esquema ... PRA QUE TUDO FUNCIONE MUITO BEM.</p>	<p>6. Desenvolvidas com sujeito parc. correferente explícito, com 6 casos (não houve dados de reduzidas).</p> <p>Ex.: [...] há que se entender os reais méritos das terapias alternativas, PARA QUE ELAS DEIXEM DE FAZER FALSAS PROMESSAS.</p>
<p>7. Reduzidas com sujeito parc. correferente explícito, com 5 casos.</p> <p>Ex.: [...] eu tenho que passar pra outras pessoas... PARA ELES PODEREM ACREDITAR, porque tem muita gente que não acredita... em milagres.</p>	<p>7. Desenvolvidas com sujeito não correferente não explícito, com 4 casos.</p> <p>Ex.: [...] ao invés de esconder, elas se mostraram reveladoras, apontando o caminho PARA QUE DESCOBRÍSSEMOS AS MEDIDAS DO ESPAÇO E DO TEMPO.</p>
<p>8. Desenvolvidas com sujeito correferente não explícito, com 1 caso.</p> <p>Ex.: [...] o ensino público ele deve ser melhorado desde a base PRA QUE SEJA tão bom quanto as escolas particulares [...]</p>	<p>8. Desenvolvidas com sujeito correferente não explícito, com 3 casos.</p> <p>Ex: Galinhas poedeiras vivem espremidas sob luz quase ininterrupta PARA QUE COMAM E BOTEM OVOS SEM PARAR.</p>
<p>9. Desenvolvidas com sujeito parc. correferente explícito, com 1 caso.</p> <p>Ex.: [...] a educadora não aceitou o meu argumento... PRA QUE MEU DESENHO FICASSE NO PADRÃO DE TODAS AS PESSOAS me fez fazer um... refazer o desenho.</p>	<p>9. Desenvolvidas com sujeito parc. correferente não explícito, com 2 casos.</p> <p>Ex.: Quanto aos outros, deveríamos parar de reproduzi-los, exceto por um pequeno número deles, que poderíamos manter em reservas PARA QUE NÃO FOSSEM EXTINTOS.</p>
	<p>10. Desenvolvidas com sujeito correferente explícito, com 1 caso (não houve dados de reduzidas).</p> <p>Ex.: PARA QUE O HOMEM PUDESSE TER CERTEZA de que o filho gerado era seu, teria que garantir que a mulher não tinha feito sexo com outros.</p>

Na convergência das duas variáveis independentes especificadas, o quadro revela que o caso menos marcado, na fala e na escrita, é de oração final reduzida com sujeito correferente não explícito. É importante salientar que se tomou como primeiro critério para a definição dessa escala apenas a frequência das ocorrências, mas pode-se perceber que pelo menos um dos

outros dois critérios arrolados por Givón (1995), a complexidade estrutural ou formal, também está presente no caso menos marcado, pois, em termos da estrutura linguística, as reduzidas contêm menos material e são menos extensas que as desenvolvidas. Em termos da complexidade cognitiva, cuja mensuração é menos visualizável, poder-se-ia chegar à mesma conclusão, já que as reduzidas demandariam um tempo de processamento linguístico menor que o das desenvolvidas. Por outro lado, isso pode ser questionado, pois se as reduzidas em geral são regidas pelo princípio da economia linguística, por codificarem aquilo que já está previsto discursivamente, os interlocutores precisariam fazer inferências para compreender o que ficou implícito no discurso, o que revelaria também certa complexidade cognitiva.

Do segundo nível da escala de marcação até o último, percebe-se que fala e escrita passam a divergir no que tange às frequências encontradas. Como aponta Givón (1995, p. 27) – cf. capítulo 3 –, essa divergência é bastante pertinente, pois a marcação linguística é dependente do contexto do qual os dados emergem. Assim, além de os dados terem sido coletados de modalidades diferentes da língua, os gêneros referentes a essas modalidades também são diferentes, o que contribui para que as frequências com que certas formas são empregadas não se coadunem, já que o emprego de uma ou de outra estrutura depende em certa medida das peculiaridades discursivas de cada gênero.

É interessante observar que, no segundo nível da escala, a fala apresenta as finais reduzidas com sujeito não correferente explícito (nível 5 na escrita), enquanto a escrita traz as desenvolvidas com essas mesmas características morfossintáticas (nível 6 na fala).

Em relação às formas mais marcadas, ou seja, aquelas mais raras em termos da frequência de uso – desprezando-se os casos em que houve categoricidade –, os dados do PORTVIX revelaram serem de baixo emprego as finais desenvolvidas com sujeito correferente não explícito (mesmo nível da escrita, na 8ª posição) e as finais desenvolvidas com sujeito parcialmente correferente explícito (nível 6 na escrita). Ambas apresentaram apenas 1 (uma) ocorrência, de maneira que a colocação desta em detrimento daquela na última posição da escala tem por base o critério da complexidade formal da estrutura, pois a última, com sujeito explícito, tem maior codificação que a primeira, teoricamente mais marcada, portanto. Já a amostra da escrita apontou como a forma mais marcada as desenvolvidas com sujeito correferente explícito, também com apenas 1 (um) caso – esse tipo de construção não ocorreu nos dados da fala.

5.6 A explicitação do sujeito como variável dependente: breves considerações

Ao final da análise de relevância estatística desta pesquisa, a variável independente que contém os fatores sujeito explícito e não explícito foi selecionada como significativamente influente na variação entre finais desenvolvidas e reduzidas nos dois *corpora*. Também ao longo da apuração dos resultados, ficou clara a correlação entre esta variável e a que diz respeito à correferencialidade ou não do sujeito da oração final. Por uma razão discursiva, semântica e estrutural, veio se entendendo que a posição da oração, a depender do gênero textual, também poderia compor esse quadro de correlação de forma considerável.

Para atestar essa relação interdependente, fez-se uma experiência que tomou a explicitação do sujeito como variável dependente, e que deslocou a anterior, finais desenvolvidas e reduzidas, para a função de variável independente, conservando-se todas as outras quatro variáveis que já haviam sido anteriormente aplicadas. Essa experiência foi feita com os dados da fala e da escrita, e o resultado encontrado esteve bem próximo do que se projetou.

A rodada com os dados da escrita trouxe resultados diferentes dos da fala, apresentando como relevantes os grupos de fatores da correferencialidade do sujeito e o da estrutura da oração – desenvolvida ou reduzida.

No entanto, com os dados do PORTVIX, a rodada de pesos relativos não podia ter sido mais exata à expectativa gerada, visto que, tendo como variável dependente a explicitação do sujeito da adverbial final, foram selecionados como estatisticamente relevantes pelo Varbrul justamente os grupos de fatores da correferencialidade do sujeito e da posição da oração. Tomando-se a variante *sujeito explícito* como a aplicação, os números foram os seguintes:

TABELA 18

Atuação das variáveis selecionadas por relevância estatística no uso de sujeito explícito em orações finais na fala

Variáveis selecionadas	Frequências	Peso relativo
Correferencialidade do sujeito da adverbial final		
<i>(a) sujeito correferente</i>	28/212=13.2%	0.36
<i>(b) sujeito parcialmente correferente</i>	6/36=16.7%	0.46
<i>(c) sujeito não correferente</i>	40/54=74.1%	0.92
Posição da adverbial final		
<i>(a) final posposta</i>	61/281=21.7%	0.46
<i>(b) final anteposta</i>	13/21=61.9%	0.87
Total	74/302=24.5%	

Fonte: PORTVIX, 2002.

Em termos de frequência relativa, verifica-se que a não correferencialidade do sujeito aumenta consideravelmente a média de sujeito explícito, de 24.5% para 74.1%, e apresenta peso relativo de 0.92, altamente favorecedor à sua ocorrência. O fator final anteposta também apresenta desvio para cima em relação à média, com 61.9%, e peso relativo de 0.87, o que indica forte favorecimento da emersão discursiva de sujeito explícito na adverbial final.

Além de confirmar a imbricação que há entre essas variáveis, outro objetivo de se fazer essa experiência e de mostrar de forma concisa esses resultados é de acenar para futuras possibilidades de pesquisa linguística com uma nova variável dependente, não apenas em orações finais, mas em outros tipos também, de forma a entender melhor o comportamento variável do sujeito das orações adverbiais, no que concerne à sua (não) explicitação.

Na próxima seção, são apresentadas as considerações finais, em termos do que se concluiu a partir de todo o desenvolvimento deste trabalho, bem como são apontadas algumas possibilidades de pesquisas futuras, aventadas ao longo do estudo feito com as orações finais na perspectiva de sua variação sintática.

CONSIDERAÇÕES FINAIS

Ao final da etapa deste trabalho em que se analisaram os dados linguísticos por meio de sua quantificação, com a finalidade de submetê-los ao tratamento de significância estatística, os resultados encontrados possibilitaram uma visão ao mesmo tempo geral e específica no que concerne à variação sintática das orações adverbiais finais na fala e na escrita, nos recortes dos gêneros textuais aqui feitos.

Nesse sentido, as ideias e hipóteses iniciais que governaram a proposição, bem como o desenvolvimento da pesquisa, puderam ser, majoritariamente, confirmadas. Uma delas foi a confirmação de que, nos *corpora* pesquisados, as finais desenvolvidas, em termos do percentual global aferido, constituem o item marcado entre as finais, e as reduzidas, o não marcado, tendo em vista as frequências de 3% e 16% encontradas, respectivamente, para as primeiras na fala (amostra PORTVIX) e na escrita (reportagens da revista Superinteressante). Aproximados esses percentuais aos encontrados nos trabalhos de Finck (2000) – 3% –, com base nos dados da fala de habitantes da cidade de Florianópolis (SC), e Azevedo (2000) – 17,2% –, a partir de dados da escrita formal de livros e jornais, ratifica-se a ideia de que a variação entre finais desenvolvidas e reduzidas apresenta um caráter amplo no sistema linguístico do português, não configurado como um fenômeno regional.

Para além da superfície das frequências globais, e para o estudo mais efetivo da variação sintática delimitada, foram controladas 8 variáveis independentes, sendo 5 de natureza linguística e 3 de caráter social, sobre as quais foi gerada a expectativa de maior influência das ligadas à estrutura interna da língua, em termos do entendimento desta área de variabilidade, a sintaxe. Em relação às variáveis sociais, não se esperava um papel importante delas na influência sobre o fenômeno aqui investigado, exceto o nível de escolarização na fala dos informantes da amostra PORTVIX. Desse modo, as duas expectativas foram confirmadas com os resultados obtidos, o que poderá ser visualizado na Tabela 19, apresentada a seguir. Nesta mesma tabela, o objetivo é trazer uma visão de conjunto dos resultados encontrados, tanto das variáveis consideradas significativamente relevantes do ponto de vista estatístico, bem como das que não foram assim definidas. Tendo como referência as finais desenvolvidas, são reapresentadas, portanto, as frequências para todas as variáveis, para os dados dos dois *corpora*, assim como os pesos relativos apenas daquelas selecionadas pelo Varbrul como significantes.

TABELA 19

Síntese dos resultados aferidos para as variáveis independentes e sua influência no uso das finais desenvolvidas nos dados da fala e da escrita

Variáveis independentes	Fala		Escrita	
	Freq.	P.R.	Freq.	P.R.
(1) Correferencialidade do sujeito da adv. final				
(a) <i>sujeito correferente</i>	1/212=0.5%	—	4/191=2.1%	0.36
(b) <i>sujeito parcialmente correferente</i>	1/36=2.8%	—	8/27=29.6%	0.81
(c) <i>sujeito não correferente</i>	7/54=13%	—	33/64=51.6%	0.76
(2) Características do sujeito da or. principal				
(a) <i>sujeito controlador/animado</i>	3/276=1.1%	0.44	18/179=10.1%	—
(b) <i>sujeito não control./inanimado/oracional</i>	6/26=23.1%	0.94	27/103=26.2%	—
(3) Voz/aspecto semântico do verbo				
(a) <i>ativa / expressão do sujeito como agente</i>	8/289=2.8%	—	37/260=14.2%	—
(b) <i>passiva / expressão do sujeito como paciente</i>	1/13=7.7%	—	8/22=36.4%	—
(4) Explicitação do sujeito na adv. final				
(a) <i>sujeito explícito</i>	8/74=10.8%	0.91	36/42=85.7%	0.97
(b) <i>sujeito não explícito</i>	1/228=0.4%	0.32	9/240=3.8%	0.35
(5) Posição da adverbial final				
(a) <i>final posposta</i>	7/281=2.5%	—	36/206=17.5%	—
(b) <i>final anteposta</i>	2/21=9.5%	—	9/76=11.8%	—
(6) Sexo/gênero				
(a) <i>masculino</i>	4/161=2.5%	—	30/199=15.1%	—
(b) <i>feminino</i>	5/141=3.5%	—	15/83=18.1%	—
(7) Faixa etária				
(a) <i>7-14 anos</i>	0/35=0%	—	—	—
(b) <i>15-25 anos</i>	4/119=3.4%	—	12/80=15%	—
(c) <i>26-49 anos</i>	5/74=6.8%	—	22/139=15.8%	—
(d) <i>50 anos ou mais</i>	0/74=0%	—	11/63=17.5%	—
(8) Nível de escolarização				
(a) <i>ensino fundamental e médio</i>	1/183=0.5%	0.25	—	—
(b) <i>ensino superior</i>	8/119=6.7%	0.84	—	—
Média global das finais desenvolvidas	9/302=3%		45/282=16%	

Fonte: Entrevistas do PORTVIX; Reportagens da Superinteressante

Confirmando o que se projeta na Tabela 19, identifica-se que a dependência da variável final desenvolvida x final reduzida, em termos estatísticos, na perspectiva de favorecimento à ocorrência de desenvolvidas e de desfavorecimento às reduzidas, encontra-se na seguinte configuração para os dados da fala: o fator sujeito não controlador com 23.1% das ocorrências, e peso relativo de 0.94; a presença de sujeito explícito na adverbial final, com 10.8% de frequência e peso relativo de 0.91; e nível de ensino superior do usuário da língua, com 6.7% de frequência relativa e peso relativo de 0.84.

Para os dados da escrita, na mesma perspectiva, os fatores sujeito parcialmente correferente, com 29.6% de frequência e 0.81 de peso relativo, e sujeito não correferente, 51.6% de uso e 0.76 de peso relativo, favorecem a emersão discursiva das finais desenvolvidas e desfavorecem a das reduzidas.

Apesar de as outras variáveis não terem sido apontadas como significativas estatisticamente, é possível entrever, analogamente, que as tendências em termos de favorecimento ou desfavorecimento a uma ou a outra variante são muito semelhantes na fala e na escrita. Nas variáveis linguísticas, somente se verificou uma inversão dessa tendência quando se analisou a posição da adverbial final, entendida como causada pela diferença entre os gêneros textuais – entrevista laboviana e reportagem – que compõem os *corpora* linguísticos desta pesquisa.

De forma abrangente, pode-se dizer que a área de variação profícua do fenômeno das orações finais, tanto na fala quanto na escrita, ocorre na explicitação do sujeito dessas orações, que decorre, sobremaneira, da sua não correferencialidade com o sujeito da oração principal, ou ainda, do que se chamou neste trabalho de correferencialidade parcial, como nos seguintes dados, da fala (1) e da escrita (2):

(1) [...] a gente tinha que descer lá em baixo [...] PRA ELES DAR PRESENÇA, se não a gente ficava com falta.

(2) Seja qual for o instrumento, a música tem sempre a finalidade de fazer o homem adormecer, PARA QUE A SEREIA POSSA MATÁ-LO.

Por isso, como se viu na parte analítica deste trabalho, compreende-se que a correlação entre a (não) correferencialidade e a (não) explicitação do sujeito da adverbial final é motivada economicamente, em que a primeira motiva a segunda, e que a maior ou menor codificação na

posição de sujeito motiva, por iconicidade, a ocorrência da final desenvolvida, no primeiro caso, e da reduzida no segundo.

Para finalizar, é importante dizer que ao longo do desenvolvimento desta dissertação vislumbraram-se algumas possibilidades futuras de pesquisa que não foram aqui feitas por uma questão metodológica e também temporal. Nesse sentido, essas potencialidades podem ser assim discriminadas:

1. Ampliar o tratamento variacionista no sentido de controlar outras variáveis que podem ser importantes para o estudo das orações finais, tais como as marcas de flexão do infinitivo pessoal e os níveis de integração entre as orações finais e as que a estas estiverem vinculadas.
2. Aplicar a metodologia de trabalho aqui adotada para o estudo e correlação com outros tipos oracionais, não apenas adverbiais, para que se tenha uma visão mais ampla do fenômeno de desenvolvimento e redução de orações, com outras configurações sintáticas, na língua portuguesa falada e escrita no Brasil.
3. Considerar, nessa ampliação do conjunto de orações, a coleta de dados numa variedade maior de gêneros textuais, com vistas a perceber se os comportamentos sintáticos variáveis das orações são dependentes dos gêneros dos quais elas emergem.
4. Fazer analogias entre a língua portuguesa e outras línguas neolatinas, como o espanhol e o francês, para que se verifiquem semelhanças e diferenças em termos da variação sintática de orações adverbiais ou não, sobretudo em função de essas línguas não apresentarem em sua estrutura de uso o infinitivo flexionado, e ver até que ponto isso pode influenciar numa menor incidência da redução de orações.
5. Tomar como variável dependente o fenômeno da (não) explicitação do sujeito em orações subordinadas, pois, como se ensaiou ao final da análise deste trabalho, essa variável tem alta correlação com a correferencialidade do sujeito, bem como com a posição em que a oração se encontra.

Essas possibilidades são elencadas com o objetivo de mostrar que o trabalho com a variação sintática no âmbito das orações e de seus constituintes é instigante e importante para o entendimento do comportamento dessas estruturas na língua portuguesa, tanto que motivou a realização desta pesquisa e apresenta terreno fértil para motivar muitas outras futuras.

REFERÊNCIAS

AZEREDO, J. C. de. **Gramática Houaiss da língua portuguesa**. 2. ed. São Paulo: Publifolha, 2008.

AZEVEDO, J. L. F. **A expressão da finalidade em português**. Rio de Janeiro: Faculdade de Letras da UFRJ, 2000. (Tese de Doutorado)

BECHARA, E. **Moderna gramática portuguesa**. 37. ed. Rio de Janeiro: Lucena, 2009.

BIBER, D. **Variation across speech and writing**. Cambridge: Cambridge University Press, 1988.

BOLINGER, D. **Aspects of Language**. New York: Harcourt Brace Jovanovich, 1975.

BRAGA, M. L. B *et al.* A ordem das orações nos discursos falados e escritos. In: RONCARATI, C.; ABRAÇADO, J. (Orgs). **Português brasileiro II: contato linguístico, heterogeneidade e história**. Niterói: EdUFF, 2008.

BRAGA, M. L. A relação semântica de tempo, o gerúndio e as orações desenvolvidas. In: ZILLES, A. M. S. (org.). **Estudos de variação linguística no Brasil e no Cone Sul**. Porto Alegre: Editora da UFRGS, 2005.

_____. Processos de redução: o caso das orações de gerúndio. In: KOCH, I. G. V. (org.) **Gramática do Português falado**. v. VI: Desenvolvimentos. Campinas: Editora da UNICAMP, 1996.

BYBEE, J.; HOPPER, P. J. Introduction to frequency and the emergence of linguistic structure. In: BYBEE, J.; HOPPER, P. J. (Ed.). **Frequency and the emergence of linguistic structure**. Amsterdam: John Benjamins, 2001, p. 1-24.

CHAFE, W. Linking intonation units in spoken English. In: HAIMAN, J.; THOMPSON, S. A. (eds.) **Clause combining in grammar and discourse**. Amsterdam: John Benjamins Publishing Company, 1988.

COLEÇÃO completa Superinteressante. São Paulo: Abril, 2005. 9 CD-ROM.

CROFT, W. **Typology and Universals**. Cambridge: Cambridge University Press, 1990.

CUNHA, C.; CINTRA, L.. **Nova gramática do português contemporâneo**. 3. ed. Rio de Janeiro: Nova Fronteira, 2001.

DIAS, N. B. **As cláusulas de finalidade**. Campinas: Instituto de Estudos da Linguagem (Unicamp), 2001. (Tese de Doutorado)

_____. Cláusulas de finalidade: relações gramaticais convergentes e divergentes na fala e na escrita. **Estudos linguísticos**, nº 34, p. 527-532.

DU BOIS, J. W. Competing motivations. In: HAIMAN, J. (ed.). **Iconicity in syntax**. Amsterdam: John Benjamins Publishing Company, 1985, p. 343-365.

FINCK, D. **Expansão e redução de cláusulas infinitivas na fala de Florianópolis**. Florianópolis: UFSC, 2000. (Dissertação de Mestrado)

FURTADO DA CUNHA, M. A. Funcionalismo. In: MARTELOTTA, M. E. (org.). **Manual de linguística**. São Paulo: Contexto, 2009.

_____. O modelo das motivações competidoras no domínio funcional da negação. **D.E.L.T.A.** 2001, vol. 17, n. 1.

FURTADO DA CUNHA, M. A.; COSTA, M. A.; CEZARIO, M. M. Pressupostos teóricos fundamentais. In: CUNHA, M. A. F da; OLIVEIRA, M. R. de; MARTELOTTA, M. E. (orgs.) **Linguística funcional teoria e prática**. Rio de Janeiro: DP&A, 2003.

GIVÓN, T. **Functionalism and grammar**. Amsterdam/Philadelphia: John Benjamins, 1995.

_____. **Syntax**. Amsterdam/Philadelphia: John Benjamins, 2001.

GÖRSKI, E. A (não) realização do sujeito e a integração de orações. **Revista Scripta**, v.5, n.9, p. 161-173. Belo Horizonte: PUC Minas, 2001.

_____. Níveis de integração de cláusulas para + infinitivo. **Estudos linguísticos**, v.29, p.88-102. Assis/SP, 2000.

GÖRSKI, E. M.; TAVARES, M. A.; FREITAG, R. M. K. Restrições de natureza cognitivo-comunicativa: marcação *versus* expressividade retórica em fenômenos variáveis. In: RONCARATI, C.; ABRAÇADO, J. (Orgs). **Português brasileiro II: contato linguístico, heterogeneidade e história**. Niterói: EdUFF, 2008.

GUY, G. R.; ZILLES, A. **Sociolinguística quantitativa: instrumental de análise**. São Paulo: Parábola Editorial, 2007.

HAIMAN, J. Iconic and economic motivation. **Language** (59). 1983, p. 781-819.

HAIMAN, J.; THOMPSON, S. 'Subordination' in Universal Grammar. **Proceedings of the Tenth Annual Meeting of Berkeley Linguistics Society**. Berkeley: Berkeley Linguistics Society, 1984.

KURY, A. da. G. **Novas lições de análise sintática**. 9. ed. São Paulo: Ática, 2008.

LABOV, W. **Padrões sociolinguísticos**. Tradução de Marcos Bagno, Maria Marta Pereira Scherre e Caroline Rodrigues Cardoso. São Paulo: Parábola Editorial, 2008.

_____. **Principles of linguistic change: social factors**. Oxford: Blackwell, 1994.

LEHMANN, C. Towards a typology of clause linkage. In: HAIMAN, J.; THOMPSON, S. (eds.). **Clause combining in grammar and discourse**. Philadelphia: John Benjamins Publishing Company, 1988.

MARTELOTTA, M. E. Funcionalismo. In: WILSON, V.; MARTELOTTA, M. E.; CEZARIO, M. M. **Linguística: fundamentos**. Rio de Janeiro: CCAA Editora, 2006.

MELLO, G. C. de. **Gramática fundamental da língua portuguesa**. Rio de Janeiro: Livraria Acadêmica, 1968.

MENEZES, V. M. C. de. **Construções infinitivas iniciadas por para**: oracionalidade e redução. Rio de Janeiro: Faculdade de Letras da UFRJ, 2001. (Tese de Doutorado)

NARO, A. J.; SCHERRE, M. M. P. Análise quantitativa e tópicos de interpretação do Varbrul. In: BRAGA, M. L.; MOLLICA, M. C. (Orgs.). **Introdução à Sociolinguística: o tratamento da variação**. São Paulo: Contexto, 2008.

NARO, A. J. Modelos quantitativos e tratamento estatístico. In: BRAGA, M. L.; MOLLICA, M. C. (orgs.). **Introdução à Sociolinguística: o tratamento da variação**. São Paulo: Contexto, 2008.

_____. Variação e funcionalidade. **Revista de estudos da linguagem**, v.7, n. 2, p. 109-120. Belo Horizonte: Faculdade de Letras, 1998.

NEVES, M. H. de M. **A gramática funcional**. São Paulo: Martins Fontes, 1997.

_____. Estudos funcionalistas no Brasil. **D.E.L.T.A.** 1999, vol. 15, n. especial, p. 70-104.

_____. **Gramática de usos do português**. São Paulo: Editora Unesp, 2000.

OLIVEIRA, M. A. de. Variável linguística: conceituação, problemas de descrição gramatical e implicações para a construção de uma teoria gramatical. **D.E.L.T.A.** 1987, vol. 3, n. 1, p. 19-34.

PAIVA, M. da. C. de. **Formas de expressão de causalidade**. Relatório de Pesquisa. Rio de Janeiro, UFRJ, 1999.

_____. A variável gênero/sexo. In: BRAGA, M. L.; MOLLICA, M. C. (Orgs.). **Introdução à Sociolinguística: o tratamento da variação**. São Paulo: Contexto, 2008.

PERINI, M. A. **Gramática descritiva do português**. São Paulo: Ática, 1995.

ROCHA LIMA, C. H. da. **Gramática normativa da língua portuguesa**. 48. ed. Rio de Janeiro: José Olympio, 2010.

SAID ALI, M. **Gramática secundária e gramática histórica da língua portuguesa**. 3. ed. Brasília: Ed. UnB, 1964.

SAUSSURE, F. de. **Curso de linguística geral**. São Paulo: Cultrix, 1975.

SANKOFF, D; SMITH, E; TAGLIAMONTE, S. **Goldvarb X**: a variable rule application for Macintosh and Windows. Department of Linguistics, University of Toronto, 2005.

SCHERRE, M. M. P. Concordância nominal e funcionalismo. **Alfa**, n. 41, p. 181-206. São Paulo, 1997.

_____. Pressupostos teóricos e suporte quantitativo. In: SILVA, G. M. de O.; SCHERRE, M. M. P. (Orgs.). **Padrões sociolinguísticos**: análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro. Rio de Janeiro: Tempo Brasileiro, 1996.

TARALLO, F. **A pesquisa sociolinguística**. São Paulo: Ática, 1985.

THOMPSON, S. A. **Grammar and written discourse**: initial vs. final purpose clauses in English. Text 5 (1-2) p.55-84, 1985.

TRASK, R. L. **Dicionário de linguagem e linguística**. Trad. Rodolfo Ilari. 2. ed. São Paulo: Contexto, 2006.

VOTRE, S. J. Relevância da variável escolaridade. In: BRAGA, M. L.; MOLLICA, M. C. (orgs.). **Introdução à Sociolinguística**: o tratamento da variação. São Paulo: Contexto, 2008.

WEINREICH, U.; LABOV, W.; HERZOG, M. **Fundamentos empíricos para uma teoria da mudança linguística**. Trad. Marcos Bagno. São Paulo: Parábola, 2006.

YACOVENCO, L. C. O projeto “O português falado na cidade de Vitória”: coleta de dados. In: LINS, M. da P. P.; YACOVENCO, L. C. (Orgs.) **Caminhos linguísticos**. Vitória: NUPLES, 2002.

ANEXOS

Anexo 1: Especificações dos símbolos para codificação dos dados no Varbrul

VARIÁVEL DEPENDENTE			
VARIANTES		SÍMBOLOS	
(a) oração subordinada adverbial final desenvolvida		D	
(b) oração subordinada adverbial final reduzida		R	
VARIÁVEIS INDEPENDENTES		FATORES	SÍMBOLOS
1. Correferencialidade do sujeito da adv. final	(a) suj. correferente		=
	(b) suj. parc. correferente		&
	(c) suj. não correferente		#
2. Características sintático-semânticas do sujeito da oração principal	(a) suj. controlador/animado		1
	(b) suj. não controlador /inanimado/oracional		2
3. Voz/aspecto semântico do verbo da adv. final	(a) ativa / verbo que expresse o sujeito agente		3
	(b) passiva / verbo que expresse o sujeito paciente		4
4. Explicitação do sujeito na adverbial final	(a) sujeito explícito		X
	(b) sujeito não explícito		W
5. Posição da adverbial final	(a) final posposta		P
	(b) final anteposta		A
6. Sexo/gênero do usuário	(a) masculino		H
	(b) feminino		M
7. Faixa etária	(a) 7 a 14 anos	U	B ³⁴
	(b) 15 a 25 anos	D	
	(c) 26 a 49 anos	T	A
	(d) 50 anos ou mais	Q	
8. Nível de escolarização	(a) fundamental	F	M
	(b) médio	I	
	(c) superior	S	

³⁴ Codificação atribuída em função da amalgamação dos fatores dos grupos *faixa etária* e *nível de escolarização* nas rodadas com os dados do PORTVIX.

Anexo 2: Rodada dos dados da fala

• GROUPS & FACTORS • 26/07/2011 19:14:49


```

-----
Group      Default  Factors
  1         R      RD
  2         =     =&#
  3         1      12
  4         3      34
  5         W      WX
  6         P      PA
  7         H      HM
  8         U      UQDT
  9         F      FIS
  
```

• CELL CREATION • 26/07/2011 19:15:03

Name of token file: Dados da fala 26 07 2.tkn
 Name of condition file: Untitled.cnd

```

(
; Identity recode: All groups included as is.
(1)
(2)
(3)
(4)
(5)
(6)
(7)
(8)
(9)
)
  
```

Number of cells: 92
 Application value(s): DR
 Total no. of factors: 20

Group		D	R	Total	%	

1 (2)		D	R			
=	N	1	211	212	70.2	sujeito correferente
	%	0.5	99.5			
&	N	1	35	36	11.9	sujeito parcialmente co
rreferente	%	2.8	97.2			
#	N	7	47	54	17.9	sujeito não correferent
e	%	13.0	87.0			
Total	N	9	293	302		
	%	3.0	97.0			

2 (3)		D	R			
1	N	3	274	277	91.7	sujeito controlador

	%	1.1	98.9			
2	N	6	19	25	8.3	sujeito não controlador
	%	24.0	76.0			
Total	N	9	293	302		
	%	3.0	97.0			
<hr/>						
3 (4)		D	R			
3	N	8	281	289	95.7	voz ativa
	%	2.8	97.2			
4	N	1	12	13	4.3	voz passiva
	%	7.7	92.3			
Total	N	9	293	302		
	%	3.0	97.0			
<hr/>						
4 (5)		D	R			
W	N	1	227	228	75.5	sujeito não explícito
	%	0.4	99.6			
X	N	8	66	74	24.5	sujeito explícito
	%	10.8	89.2			
Total	N	9	293	302		
	%	3.0	97.0			
<hr/>						
5 (6)		D	R			
P	N	7	274	281	93.0	final posposta
	%	2.5	97.5			
A	N	2	19	21	7.0	final anteposta
	%	9.5	90.5			
Total	N	9	293	302		
	%	3.0	97.0			
<hr/>						
6 (7)		D	R			
H	N	4	157	161	53.3	homens
	%	2.5	97.5			
M	N	5	136	141	46.7	mulheres
	%	3.5	96.5			
Total	N	9	293	302		
	%	3.0	97.0			
<hr/>						
7 (8)		D	R			
U	N	0	35	35	11.6	
	%	0.0	100.0		* KnockOut *	7-14 anos
Q	N	0	74	74	24.5	
	%	0.0	100.0		* KnockOut *	50 ou mais
D	N	4	115	119	39.4	15-25 anos
	%	3.4	96.6			
T	N	5	69	74	24.5	26-49 anos
	%	6.8	93.2			

Total N	9	293	302		
%	3.0	97.0			
<hr/>					
8 (9)	D	R			
F N	0	90	90	29.8	ensino fundamental
%	0.0	100.0		* KnockOut *	
I N	1	92	93	30.8	ensino médio
%	1.1	98.9			
S N	8	111	119	39.4	ensino superior
%	6.7	93.3			
Total N	9	293	302		
%	3.0	97.0			
<hr/>					
TOTAL N	9	293	302		
%	3.0	97.0			

Name of new cell file: .cel

• CELL CREATION • 26/07/2011 19:21:50 •.....
 •.....

Name of token file: Dados da fala 26 07 2.tkn
 Name of condition file: Untitled.cnd

(
 ; Identity recode: All groups included as is.

(1)
 (2)
 (3)
 (4)
 (5)
 (6)
 (7)

;amalgamando faixa etária de 7-14 com 15-25, e 16-49 com 50 ou +

(8 (B (col 8 U))
 (B (col 8 D))
 (A (col 8 T))
 (A (col 8 Q)))

;amalgamando nível de escolarização fundamental com médio

(9 (M (col 9 F))
 (M (col 9 I)))
)

Number of cells: 67
 Application value(s): DR
 Total no. of factors: 17

Group	D	R	Total	%	
<hr/>					
1 (2)	D	R			
= N	1	211	212	70.2	sujeito correferente
%	0.5	99.5			
& N	1	35	36	11.9	sujeito parcialmente co
rreferente					
%	2.8	97.2			

e	#	N	7	47	54	17.9	sujeito não correferent
		%	13.0	87.0			
Total	N		9	293	302		
	%		3.0	97.0			
<hr/>							
2	(3)		D	R			
1	N		3	274	277	91.7	sujeito controlador
	%		1.1	98.9			
2	N		6	19	25	8.3	sujeito não controlador
	%		24.0	76.0			
Total	N		9	293	302		
	%		3.0	97.0			
<hr/>							
3	(4)		D	R			
3	N		8	281	289	95.7	voz ativa
	%		2.8	97.2			
4	N		1	12	13	4.3	voz passiva
	%		7.7	92.3			
Total	N		9	293	302		
	%		3.0	97.0			
<hr/>							
4	(5)		D	R			
W	N		1	227	228	75.5	sujeito não explícito
	%		0.4	99.6			
X	N		8	66	74	24.5	sujeito explícito
	%		10.8	89.2			
Total	N		9	293	302		
	%		3.0	97.0			
<hr/>							
5	(6)		D	R			
P	N		7	274	281	93.0	final posposta
	%		2.5	97.5			
A	N		2	19	21	7.0	final anteposta
	%		9.5	90.5			
Total	N		9	293	302		
	%		3.0	97.0			
<hr/>							
6	(7)		D	R			
H	N		4	157	161	53.3	homens
	%		2.5	97.5			
M	N		5	136	141	46.7	mulheres
	%		3.5	96.5			
Total	N		9	293	302		
	%		3.0	97.0			
<hr/>							
7	(8)		D	R			
B	N		4	150	154	51.0	7-14 e 15-25 anos
	%		2.6	97.4			

A	N	5	143	148	49.0	26-49 e 50 anos ou mais
	%	3.4	96.6			
Total	N	9	293	302		
	%	3.0	97.0			

8 (9)		D	R			
M	N	1	182	183	60.6	ensino fundamental e mé
dio	%	0.5	99.5			
S	N	8	111	119	39.4	ensino superior
	%	6.7	93.3			
Total	N	9	293	302		
	%	3.0	97.0			

TOTAL	N	9	293	302		
	%	3.0	97.0			

Rodada de pesos relativos após amalgamações necessárias

```

Name of new cell file: .cel
• BINOMIAL VARBRUL • 26/07/2011 19:22:17 .....
.....
Name of cell file: .cel
Averaging by weighting factors.
Threshold, step-up/down: 0.050001
Stepping up...

----- Level # 0 -----

Run # 1, 1 cells:
Convergence at Iteration 2
Input 0.030
Log likelihood = -40.483

----- Level # 1 -----

Run # 2, 3 cells:
Convergence at Iteration 6
Input 0.011
Group # 1 -- =: 0.304, &: 0.724, #: 0.932
Log likelihood = -31.751 Significance = 0.000

Run # 3, 2 cells:
Convergence at Iteration 5
Input 0.014
Group # 2 -- 1: 0.431, 2: 0.956
Log likelihood = -30.337 Significance = 0.000

```

Run # 4, 2 cells:
Convergence at Iteration 4
Input 0.029
Group # 3 -- 3: 0.488, 4: 0.736
Log likelihood = -40.110 Significance = 0.405

Run # 5, 2 cells:
Convergence at Iteration 6
Input 0.010
Group # 4 -- W: 0.308, X: 0.924
Log likelihood = -31.775 Significance = 0.000

Run # 6, 2 cells:
Convergence at Iteration 5
Input 0.027
Group # 5 -- P: 0.475, A: 0.789
Log likelihood = -39.364 Significance = 0.144

Run # 7, 2 cells:
Convergence at Iteration 4
Input 0.029
Group # 6 -- H: 0.458, M: 0.548
Log likelihood = -40.337 Significance = 0.606

Run # 8, 2 cells:
Convergence at Iteration 4
Input 0.030
Group # 7 -- B: 0.467, A: 0.534
Log likelihood = -40.404 Significance = 0.693

Run # 9, 2 cells:
Convergence at Iteration 5
Input 0.015
Group # 8 -- M: 0.266, S: 0.826
Log likelihood = -35.529 Significance = 0.003

Add Group # 2 with factors 12

----- Level # 2 -----

Run # 10, 6 cells:
Convergence at Iteration 8
Input 0.008
Group # 1 -- =: 0.340, &: 0.770, #: 0.859
Group # 2 -- 1: 0.450, 2: 0.903
Log likelihood = -27.208 Significance = 0.046

Run # 11, 4 cells:
Convergence at Iteration 5
Input 0.014
Group # 2 -- 1: 0.431, 2: 0.955
Group # 3 -- 3: 0.497, 4: 0.566
Log likelihood = -30.313 Significance = 0.833

Run # 12, 4 cells:
Convergence at Iteration 6
Input 0.006
Group # 2 -- 1: 0.439, 2: 0.939
Group # 4 -- W: 0.326, X: 0.904

Log likelihood = -24.405 Significance = 0.001

Run # 13, 4 cells:

Convergence at Iteration 6

Input 0.013

Group # 2 -- 1: 0.429, 2: 0.959

Group # 5 -- P: 0.471, A: 0.826

Log likelihood = -29.118 Significance = 0.125

Run # 14, 4 cells:

Convergence at Iteration 5

Input 0.014

Group # 2 -- 1: 0.430, 2: 0.957

Group # 6 -- H: 0.441, M: 0.567

Log likelihood = -30.094 Significance = 0.490

Run # 15, 4 cells:

Convergence at Iteration 5

Input 0.014

Group # 2 -- 1: 0.431, 2: 0.957

Group # 7 -- B: 0.458, A: 0.544

Log likelihood = -30.225 Significance = 0.651

Run # 16, 4 cells:

Convergence at Iteration 6

Input 0.007

Group # 2 -- 1: 0.430, 2: 0.959

Group # 8 -- M: 0.261, S: 0.833

Log likelihood = -25.764 Significance = 0.005

Add Group # 4 with factors WX

----- Level # 3 -----

Run # 17, 11 cells:

Convergence at Iteration 10

Input 0.004

Group # 1 -- =: 0.394, &: 0.821, #: 0.662

Group # 2 -- 1: 0.441, 2: 0.932

Group # 4 -- W: 0.337, X: 0.889

Log likelihood = -23.470 Significance = 0.405

Run # 18, 8 cells:

Convergence at Iteration 7

Input 0.006

Group # 2 -- 1: 0.440, 2: 0.937

Group # 3 -- 3: 0.491, 4: 0.685

Group # 4 -- W: 0.322, X: 0.908

Log likelihood = -24.251 Significance = 0.595

Run # 19, 8 cells:

Convergence at Iteration 6

Input 0.005

Group # 2 -- 1: 0.434, 2: 0.949

Group # 4 -- W: 0.331, X: 0.898

Group # 5 -- P: 0.479, A: 0.756

Log likelihood = -23.762 Significance = 0.263

Run # 20, 8 cells:

Convergence at Iteration 6
Input 0.006
Group # 2 -- 1: 0.438, 2: 0.940
Group # 4 -- W: 0.327, X: 0.903
Group # 6 -- H: 0.457, M: 0.549
Log likelihood = -24.299 Significance = 0.658

Run # 21, 8 cells:
Convergence at Iteration 6
Input 0.006
Group # 2 -- 1: 0.439, 2: 0.939
Group # 4 -- W: 0.326, X: 0.904
Group # 7 -- B: 0.491, A: 0.509
Log likelihood = -24.401 Significance = 0.933

Run # 22, 8 cells:
Convergence at Iteration 7
Input 0.003
Group # 2 -- 1: 0.437, 2: 0.942
Group # 4 -- W: 0.322, X: 0.908
Group # 8 -- M: 0.255, S: 0.838
Log likelihood = -20.283 Significance = 0.007

Add Group # 8 with factors MS

----- Level # 4 -----

Run # 23, 20 cells:
Convergence at Iteration 12
Input 0.002
Group # 1 -- =: 0.402, &: 0.855, #: 0.595
Group # 2 -- 1: 0.436, 2: 0.944
Group # 4 -- W: 0.333, X: 0.895
Group # 8 -- M: 0.253, S: 0.841
Log likelihood = -19.359 Significance = 0.410

Run # 24, 13 cells:
Convergence at Iteration 8
Input 0.003
Group # 2 -- 1: 0.437, 2: 0.942
Group # 3 -- 3: 0.497, 4: 0.560
Group # 4 -- W: 0.323, X: 0.907
Group # 8 -- M: 0.257, S: 0.837
Log likelihood = -20.271 Significance = 0.883

Run # 25, 14 cells:
Convergence at Iteration 8
Input 0.002
Group # 2 -- 1: 0.433, 2: 0.952
Group # 4 -- W: 0.330, X: 0.898
Group # 5 -- P: 0.477, A: 0.774
Group # 8 -- M: 0.257, S: 0.836
Log likelihood = -19.695 Significance = 0.282

Run # 26, 16 cells:
Convergence at Iteration 9
Input 0.002
Group # 2 -- 1: 0.437, 2: 0.944
Group # 4 -- W: 0.320, X: 0.911

Group # 6 -- H: 0.535, M: 0.460
Group # 8 -- M: 0.248, S: 0.846
Log likelihood = -20.225 Significance = 0.739

Run # 27, 16 cells:

Convergence at Iteration 16

Input 0.001

Group # 2 -- 1: 0.431, 2: 0.956

Group # 4 -- W: 0.298, X: 0.933

Group # 7 -- B: 0.699, A: 0.293

Group # 8 -- M: 0.185, S: 0.907

Log likelihood = -18.867 Significance = 0.094

No remaining groups significant

Groups selected while stepping up:

2 (características sintático-semânticas do sujeito da oração principal
)

4 (explicitação do sujeito da adverbial final)

8 (nível de escolarização)

Best stepping up run: #22

Stepping down...

----- Level # 8 -----

Run # 28, 67 cells:

No Convergence at Iteration 20

Input 0.001

Group # 1 -- =: 0.342, &: 0.832, #: 0.819

Group # 2 -- 1: 0.432, 2: 0.954

Group # 3 -- 3: 0.497, 4: 0.569

Group # 4 -- W: 0.344, X: 0.880

Group # 5 -- P: 0.483, A: 0.717

Group # 6 -- H: 0.294, M: 0.732

Group # 7 -- B: 0.889, A: 0.103

Group # 8 -- M: 0.132, S: 0.948

Log likelihood = -16.038

----- Level # 7 -----

Run # 29, 40 cells:

No Convergence at Iteration 20

Input 0.001

Group # 2 -- 1: 0.423, 2: 0.969

Group # 3 -- 3: 0.493, 4: 0.657

Group # 4 -- W: 0.300, X: 0.931

Group # 5 -- P: 0.476, A: 0.786

Group # 6 -- H: 0.295, M: 0.730

Group # 7 -- B: 0.849, A: 0.143

Group # 8 -- M: 0.161, S: 0.927

Log likelihood = -17.286 Significance = 0.289

Run # 30, 52 cells:

No Convergence at Iteration 20

Input 0.001
Group # 1 -- =: 0.300, &: 0.784, #: 0.922
Group # 3 -- 3: 0.498, 4: 0.534
Group # 4 -- W: 0.375, X: 0.828
Group # 5 -- P: 0.490, A: 0.632
Group # 6 -- H: 0.337, M: 0.684
Group # 7 -- B: 0.845, A: 0.146
Group # 8 -- M: 0.147, S: 0.937
Log likelihood = -21.308 Significance = 0.002

Run # 31, 58 cells:
No Convergence at Iteration 20
Input 0.001
Group # 1 -- =: 0.340, &: 0.831, #: 0.823
Group # 2 -- 1: 0.432, 2: 0.954
Group # 4 -- W: 0.344, X: 0.879
Group # 5 -- P: 0.483, A: 0.714
Group # 6 -- H: 0.291, M: 0.734
Group # 7 -- B: 0.888, A: 0.104
Group # 8 -- M: 0.132, S: 0.948
Log likelihood = -16.052 Significance = 0.873

Run # 32, 48 cells:
No Convergence at Iteration 20
Input 0.001
Group # 1 -- =: 0.256, &: 0.859, #: 0.952
Group # 2 -- 1: 0.441, 2: 0.934
Group # 3 -- 3: 0.500, 4: 0.506
Group # 5 -- P: 0.480, A: 0.745
Group # 6 -- H: 0.287, M: 0.738
Group # 7 -- B: 0.901, A: 0.091
Group # 8 -- M: 0.111, S: 0.961
Log likelihood = -18.255 Significance = 0.039

Run # 33, 54 cells:
No Convergence at Iteration 20
Input 0.001
Group # 1 -- =: 0.329, &: 0.868, #: 0.825
Group # 2 -- 1: 0.433, 2: 0.952
Group # 3 -- 3: 0.498, 4: 0.546
Group # 4 -- W: 0.336, X: 0.890
Group # 6 -- H: 0.283, M: 0.743
Group # 7 -- B: 0.895, A: 0.097
Group # 8 -- M: 0.128, S: 0.951
Log likelihood = -16.285 Significance = 0.487

Run # 34, 55 cells:
No Convergence at Iteration 20
Input 0.001
Group # 1 -- =: 0.345, &: 0.847, #: 0.797
Group # 2 -- 1: 0.434, 2: 0.950
Group # 3 -- 3: 0.494, 4: 0.634
Group # 4 -- W: 0.343, X: 0.881
Group # 5 -- P: 0.481, A: 0.732
Group # 7 -- B: 0.782, A: 0.209
Group # 8 -- M: 0.155, S: 0.931
Log likelihood = -16.868 Significance = 0.198

Run # 35, 57 cells:
 Convergence at Iteration 12
 Input 0.002
 Group # 1 -- =: 0.406, &: 0.838, #: 0.596
 Group # 2 -- 1: 0.433, 2: 0.951
 Group # 3 -- 3: 0.495, 4: 0.604
 Group # 4 -- W: 0.341, X: 0.884
 Group # 5 -- P: 0.480, A: 0.743
 Group # 6 -- H: 0.550, M: 0.443
 Group # 8 -- M: 0.253, S: 0.840
 Log likelihood = -18.820 Significance = 0.019

Run # 36, 49 cells:
 Convergence at Iteration 11
 Input 0.003
 Group # 1 -- =: 0.388, &: 0.833, #: 0.673
 Group # 2 -- 1: 0.438, 2: 0.940
 Group # 3 -- 3: 0.495, 4: 0.616
 Group # 4 -- W: 0.333, X: 0.895
 Group # 5 -- P: 0.477, A: 0.774
 Group # 6 -- H: 0.450, M: 0.557
 Group # 7 -- B: 0.553, A: 0.445
 Log likelihood = -22.534 Significance = 0.000

Cut Group # 3 with factors 34

----- Level # 6 -----

Run # 37, 34 cells:
 No Convergence at Iteration 20
 Input 0.001
 Group # 2 -- 1: 0.423, 2: 0.969
 Group # 4 -- W: 0.300, X: 0.932
 Group # 5 -- P: 0.476, A: 0.779
 Group # 6 -- H: 0.289, M: 0.736
 Group # 7 -- B: 0.847, A: 0.145
 Group # 8 -- M: 0.161, S: 0.927
 Log likelihood = -17.354 Significance = 0.276

Run # 38, 44 cells:
 No Convergence at Iteration 20
 Input 0.001
 Group # 1 -- =: 0.299, &: 0.785, #: 0.923
 Group # 4 -- W: 0.375, X: 0.829
 Group # 5 -- P: 0.490, A: 0.627
 Group # 6 -- H: 0.335, M: 0.687
 Group # 7 -- B: 0.845, A: 0.146
 Group # 8 -- M: 0.146, S: 0.938
 Log likelihood = -21.313 Significance = 0.002

Run # 39, 39 cells:
 No Convergence at Iteration 20
 Input 0.001
 Group # 1 -- =: 0.256, &: 0.859, #: 0.952
 Group # 2 -- 1: 0.441, 2: 0.934
 Group # 5 -- P: 0.480, A: 0.745
 Group # 6 -- H: 0.287, M: 0.738
 Group # 7 -- B: 0.901, A: 0.092

Group # 8 -- M: 0.112, S: 0.960
Log likelihood = -18.256 Significance = 0.039

Run # 40, 45 cells:

No Convergence at Iteration 20

Input 0.001

Group # 1 -- =: 0.328, &: 0.867, #: 0.827

Group # 2 -- 1: 0.433, 2: 0.951

Group # 4 -- W: 0.337, X: 0.890

Group # 6 -- H: 0.282, M: 0.745

Group # 7 -- B: 0.895, A: 0.097

Group # 8 -- M: 0.127, S: 0.951

Log likelihood = -16.291 Significance = 0.493

Run # 41, 46 cells:

No Convergence at Iteration 20

Input 0.001

Group # 1 -- =: 0.342, &: 0.847, #: 0.806

Group # 2 -- 1: 0.434, 2: 0.950

Group # 4 -- W: 0.346, X: 0.877

Group # 5 -- P: 0.481, A: 0.731

Group # 7 -- B: 0.778, A: 0.213

Group # 8 -- M: 0.157, S: 0.930

Log likelihood = -16.933 Significance = 0.189

Run # 42, 48 cells:

Convergence at Iteration 12

Input 0.002

Group # 1 -- =: 0.404, &: 0.839, #: 0.607

Group # 2 -- 1: 0.434, 2: 0.951

Group # 4 -- W: 0.342, X: 0.882

Group # 5 -- P: 0.480, A: 0.740

Group # 6 -- H: 0.544, M: 0.450

Group # 8 -- M: 0.254, S: 0.840

Log likelihood = -18.856 Significance = 0.018

Run # 43, 42 cells:

Convergence at Iteration 11

Input 0.003

Group # 1 -- =: 0.384, &: 0.832, #: 0.686

Group # 2 -- 1: 0.438, 2: 0.941

Group # 4 -- W: 0.339, X: 0.886

Group # 5 -- P: 0.476, A: 0.780

Group # 6 -- H: 0.449, M: 0.559

Group # 7 -- B: 0.551, A: 0.447

Log likelihood = -22.592 Significance = 0.000

Cut Group # 5 with factors PA

----- Level # 5 -----

Run # 44, 23 cells:

No Convergence at Iteration 20

Input 0.001

Group # 2 -- 1: 0.427, 2: 0.963

Group # 4 -- W: 0.293, X: 0.938

Group # 6 -- H: 0.267, M: 0.760

Group # 7 -- B: 0.855, A: 0.136

Group # 8 -- M: 0.162, S: 0.926

Log likelihood = -17.894 Significance = 0.202

Run # 45, 32 cells:

No Convergence at Iteration 20

Input 0.001

Group # 1 -- =: 0.293, &: 0.810, #: 0.924

Group # 4 -- W: 0.371, X: 0.836

Group # 6 -- H: 0.330, M: 0.692

Group # 7 -- B: 0.849, A: 0.142

Group # 8 -- M: 0.145, S: 0.939

Log likelihood = -21.444 Significance = 0.002

Run # 46, 28 cells:

No Convergence at Iteration 20

Input 0.001

Group # 1 -- =: 0.241, &: 0.882, #: 0.959

Group # 2 -- 1: 0.443, 2: 0.927

Group # 6 -- H: 0.283, M: 0.743

Group # 7 -- B: 0.906, A: 0.087

Group # 8 -- M: 0.110, S: 0.961

Log likelihood = -18.726 Significance = 0.030

Run # 47, 35 cells:

No Convergence at Iteration 20

Input 0.001

Group # 1 -- =: 0.328, &: 0.884, #: 0.811

Group # 2 -- 1: 0.435, 2: 0.947

Group # 4 -- W: 0.339, X: 0.886

Group # 7 -- B: 0.780, A: 0.212

Group # 8 -- M: 0.151, S: 0.934

Log likelihood = -17.239 Significance = 0.176

Run # 48, 36 cells:

Convergence at Iteration 12

Input 0.002

Group # 1 -- =: 0.396, &: 0.861, #: 0.606

Group # 2 -- 1: 0.436, 2: 0.946

Group # 4 -- W: 0.332, X: 0.896

Group # 6 -- H: 0.549, M: 0.444

Group # 8 -- M: 0.244, S: 0.850

Log likelihood = -19.253 Significance = 0.016

Run # 49, 30 cells:

Convergence at Iteration 11

Input 0.004

Group # 1 -- =: 0.395, &: 0.813, #: 0.667

Group # 2 -- 1: 0.441, 2: 0.932

Group # 4 -- W: 0.335, X: 0.892

Group # 6 -- H: 0.432, M: 0.577

Group # 7 -- B: 0.564, A: 0.434

Log likelihood = -23.338 Significance = 0.000

Cut Group # 1 with factors =&#

----- Level # 4 -----

Run # 50, 12 cells:

No Convergence at Iteration 20

Input 0.003

Group # 4 -- W: 0.297, X: 0.934
Group # 6 -- H: 0.335, M: 0.687
Group # 7 -- B: 0.745, A: 0.247
Group # 8 -- M: 0.202, S: 0.892
Log likelihood = -25.444 Significance = 0.000

Run # 51, 12 cells:
No Convergence at Iteration 20
Input 0.005
Group # 2 -- 1: 0.427, 2: 0.963
Group # 6 -- H: 0.307, M: 0.717
Group # 7 -- B: 0.756, A: 0.236
Group # 8 -- M: 0.207, S: 0.887
Log likelihood = -24.454 Significance = 0.000

Run # 52, 16 cells:
Convergence at Iteration 16
Input 0.001
Group # 2 -- 1: 0.431, 2: 0.956
Group # 4 -- W: 0.298, X: 0.933
Group # 7 -- B: 0.699, A: 0.293
Group # 8 -- M: 0.185, S: 0.907
Log likelihood = -18.867 Significance = 0.171

Run # 53, 16 cells:
Convergence at Iteration 9
Input 0.002
Group # 2 -- 1: 0.437, 2: 0.944
Group # 4 -- W: 0.320, X: 0.911
Group # 6 -- H: 0.535, M: 0.460
Group # 8 -- M: 0.248, S: 0.846
Log likelihood = -20.225 Significance = 0.034

Run # 54, 15 cells:
Convergence at Iteration 12
Input 0.005
Group # 2 -- 1: 0.438, 2: 0.942
Group # 4 -- W: 0.325, X: 0.905
Group # 6 -- H: 0.416, M: 0.596
Group # 7 -- B: 0.559, A: 0.439
Log likelihood = -24.216 Significance = 0.000

Cut Group # 6 with factors HM

----- Level # 3 -----

Run # 55, 8 cells:
Convergence at Iteration 9
Input 0.004
Group # 4 -- W: 0.297, X: 0.934
Group # 7 -- B: 0.617, A: 0.378
Group # 8 -- M: 0.228, S: 0.867
Log likelihood = -26.185 Significance = 0.000

Run # 56, 8 cells:
Convergence at Iteration 8
Input 0.006
Group # 2 -- 1: 0.428, 2: 0.962
Group # 7 -- B: 0.600, A: 0.396

Group # 8 -- M: 0.231, S: 0.864
Log likelihood = -25.300 Significance = 0.000

Run # 57, 8 cells:
Convergence at Iteration 7
Input 0.003
Group # 2 -- 1: 0.437, 2: 0.942
Group # 4 -- W: 0.322, X: 0.908
Group # 8 -- M: 0.255, S: 0.838
Log likelihood = -20.283 Significance = 0.094

Run # 58, 8 cells:
Convergence at Iteration 6
Input 0.006
Group # 2 -- 1: 0.439, 2: 0.939
Group # 4 -- W: 0.326, X: 0.904
Group # 7 -- B: 0.491, A: 0.509
Log likelihood = -24.401 Significance = 0.001

Cut Group # 7 with factors BA

----- Level # 2 -----

Run # 59, 4 cells:
Convergence at Iteration 6
Input 0.005
Group # 4 -- W: 0.305, X: 0.927
Group # 8 -- M: 0.262, S: 0.831
Log likelihood = -26.896 Significance = 0.000

Run # 60, 4 cells:
Convergence at Iteration 6
Input 0.007
Group # 2 -- 1: 0.430, 2: 0.959
Group # 8 -- M: 0.261, S: 0.833
Log likelihood = -25.764 Significance = 0.001

Run # 61, 4 cells:
Convergence at Iteration 6
Input 0.006
Group # 2 -- 1: 0.439, 2: 0.939
Group # 4 -- W: 0.326, X: 0.904
Log likelihood = -24.405 Significance = 0.007

All remaining groups significant

Groups eliminated while stepping down: 3 5 1 6 7
Best stepping up run: #22
Best stepping down run: #57

Anexo 3: Rodada da fala sem os dados dos informantes de 7 a 14 anos e de 50 anos ou mais

• CELL CREATION • 26/07/2011 19:27:53 •

Name of token file: Dados da fala 26 07 2.tkn
Name of condition file: Condições fala 26 07.cnd

(
; Identity recode: All groups included as is.

(1)

(2)

(3)

(4)

(5)

(6)

(7)

(8(nil (col 8 U))

(nil (col 8 Q)))

(9(M (col 9 F))

(M (col 9 I)))

)

Number of cells: 56

Application value(s): DR

Total no. of factors: 17

Group		D	R	Total	%	

1 (2)		D	R			
= N		1	129	130	67.4	sujeito correferente
%		0.8	99.2			
& N		1	22	23	11.9	sujeito parcialmente co
rreferente						
%		4.3	95.7			
# N		7	33	40	20.7	sujeito não correferent
e						
%		17.5	82.5			
Total N		9	184	193		
%		4.7	95.3			

2 (3)		D	R			
1 N		3	168	171	88.6	sujeito controlador
%		1.8	98.2			
2 N		6	16	22	11.4	sujeito não controlador
%		27.3	72.7			
Total N		9	184	193		
%		4.7	95.3			

3 (4)		D	R			
3 N		8	174	182	94.3	voz ativa
%		4.4	95.6			
4 N		1	10	11	5.7	voz passiva

		%	9.1	90.9		
Total	N		9	184	193	
	%		4.7	95.3		
<hr/>						
4 (5)			D	R		
W	N		1	137	138	71.5
	%		0.7	99.3		sujeito não explícito
X	N		8	47	55	28.5
	%		14.5	85.5		sujeito explícito
Total	N		9	184	193	
	%		4.7	95.3		
<hr/>						
5 (6)			D	R		
P	N		7	171	178	92.2
	%		3.9	96.1		final posposta
A	N		2	13	15	7.8
	%		13.3	86.7		final anteposta
Total	N		9	184	193	
	%		4.7	95.3		
<hr/>						
6 (7)			D	R		
H	N		4	102	106	54.9
	%		3.8	96.2		homens
M	N		5	82	87	45.1
	%		5.7	94.3		mulheres
Total	N		9	184	193	
	%		4.7	95.3		
<hr/>						
7 (8)			D	R		
D	N		4	115	119	61.7
	%		3.4	96.6		15-25 anos
T	N		5	69	74	38.3
	%		6.8	93.2		26-49 anos
Total	N		9	184	193	
	%		4.7	95.3		
<hr/>						
8 (9)			D	R		
M	N		1	126	127	65.8
	%		0.8	99.2		ensino fundamental e médio
S	N		8	58	66	34.2
	%		12.1	87.9		ensino superior
Total	N		9	184	193	
	%		4.7	95.3		
<hr/>						
TOTAL	N		9	184	193	
	%		4.7	95.3		

Name of new cell file: Condições fala 26 07.cel

• BINOMIAL VARBRUL • 26/07/2011 19:28:31
.....

Name of cell file: Condições fala 26 07.cel

Averaging by weighting factors.

Threshold, step-up/down: 0.050001

Stepping up...

----- Level # 0 -----

Run # 1, 1 cells:

Convergence at Iteration 2

Input 0.047

Log likelihood = -36.376

----- Level # 1 -----

Run # 2, 3 cells:

Convergence at Iteration 6

Input 0.019

Group # 1 -- =: 0.290, &: 0.705, #: 0.918

Log likelihood = -28.526 Significance = 0.000

Run # 3, 2 cells:

Convergence at Iteration 5

Input 0.025

Group # 2 -- 1: 0.414, 2: 0.937

Log likelihood = -27.994 Significance = 0.000

Run # 4, 2 cells:

Convergence at Iteration 4

Input 0.046

Group # 3 -- 3: 0.489, 4: 0.675

Log likelihood = -36.169 Significance = 0.527

Run # 5, 2 cells:

Convergence at Iteration 6

Input 0.018

Group # 4 -- W: 0.290, X: 0.905

Log likelihood = -28.734 Significance = 0.000

Run # 6, 2 cells:

Convergence at Iteration 5

Input 0.043

Group # 5 -- P: 0.474, A: 0.772

Log likelihood = -35.402 Significance = 0.171

Run # 7, 2 cells:

Convergence at Iteration 4

Input 0.046

Group # 6 -- H: 0.451, M: 0.560

Log likelihood = -36.168 Significance = 0.525

Run # 8, 2 cells:

Convergence at Iteration 4

Input 0.044

Group # 7 -- D: 0.430, T: 0.611
Log likelihood = -35.804 Significance = 0.288

Run # 9, 2 cells:
Convergence at Iteration 6
Input 0.021
Group # 8 -- M: 0.274, S: 0.867
Log likelihood = -30.216 Significance = 0.000

Add Group # 2 with factors 12

----- Level # 2 -----

Run # 10, 6 cells:
Convergence at Iteration 8
Input 0.015
Group # 1 -- =: 0.327, &: 0.751, #: 0.847
Group # 2 -- 1: 0.438, 2: 0.873
Log likelihood = -24.988 Significance = 0.050

Run # 11, 4 cells:
Convergence at Iteration 5
Input 0.025
Group # 2 -- 1: 0.414, 2: 0.936
Group # 3 -- 3: 0.499, 4: 0.515
Log likelihood = -27.992 Significance = 0.962

Run # 12, 4 cells:
Convergence at Iteration 6
Input 0.011
Group # 2 -- 1: 0.422, 2: 0.920
Group # 4 -- W: 0.305, X: 0.888
Log likelihood = -22.534 Significance = 0.001

Run # 13, 4 cells:
Convergence at Iteration 6
Input 0.022
Group # 2 -- 1: 0.412, 2: 0.940
Group # 5 -- P: 0.470, A: 0.806
Log likelihood = -26.961 Significance = 0.160

Run # 14, 4 cells:
Convergence at Iteration 5
Input 0.024
Group # 2 -- 1: 0.414, 2: 0.938
Group # 6 -- H: 0.443, M: 0.569
Log likelihood = -27.760 Significance = 0.496

Run # 15, 4 cells:
Convergence at Iteration 5
Input 0.023
Group # 2 -- 1: 0.414, 2: 0.938
Group # 7 -- D: 0.426, T: 0.618
Log likelihood = -27.441 Significance = 0.295

Run # 16, 4 cells:
Convergence at Iteration 7
Input 0.009
Group # 2 -- 1: 0.407, 2: 0.950

Group # 8 -- M: 0.255, S: 0.887
Log likelihood = -21.780 Significance = 0.000

Add Group # 8 with factors MS

----- Level # 3 -----

Run # 17, 11 cells:

Convergence at Iteration 9

Input 0.005

Group # 1 -- =: 0.296, &: 0.848, #: 0.862

Group # 2 -- 1: 0.431, 2: 0.896

Group # 8 -- M: 0.239, S: 0.902

Log likelihood = -18.304 Significance = 0.034

Run # 18, 8 cells:

Convergence at Iteration 7

Input 0.009

Group # 2 -- 1: 0.407, 2: 0.948

Group # 3 -- 3: 0.494, 4: 0.602

Group # 8 -- M: 0.254, S: 0.889

Log likelihood = -21.736 Significance = 0.774

Run # 19, 8 cells:

Convergence at Iteration 8

Input 0.004

Group # 2 -- 1: 0.417, 2: 0.931

Group # 4 -- W: 0.295, X: 0.899

Group # 8 -- M: 0.247, S: 0.896

Log likelihood = -16.817 Significance = 0.003

Run # 20, 8 cells:

Convergence at Iteration 8

Input 0.008

Group # 2 -- 1: 0.403, 2: 0.954

Group # 5 -- P: 0.464, A: 0.848

Group # 8 -- M: 0.247, S: 0.895

Log likelihood = -20.615 Significance = 0.135

Run # 21, 8 cells:

Convergence at Iteration 7

Input 0.009

Group # 2 -- 1: 0.407, 2: 0.950

Group # 6 -- H: 0.500, M: 0.500

Group # 8 -- M: 0.255, S: 0.887

*** Warning, negative change in likelihood (-0.00001416) replaced by 0
.0.

Log likelihood = -21.780 Significance = 1.000

Run # 22, 8 cells:

Convergence at Iteration 8

Input 0.009

Group # 2 -- 1: 0.407, 2: 0.949

Group # 7 -- D: 0.494, T: 0.510

Group # 8 -- M: 0.257, S: 0.886

Log likelihood = -21.777 Significance = 0.947

Add Group # 4 with factors WX

----- Level # 4 -----

Run # 23, 20 cells:

Convergence at Iteration 11

Input 0.003

Group # 1 -- =: 0.380, &: 0.814, #: 0.678

Group # 2 -- 1: 0.423, 2: 0.919

Group # 4 -- W: 0.328, X: 0.858

Group # 8 -- M: 0.246, S: 0.896

Log likelihood = -15.968 Significance = 0.439

Run # 24, 13 cells:

Convergence at Iteration 9

Input 0.004

Group # 2 -- 1: 0.417, 2: 0.931

Group # 3 -- 3: 0.500, 4: 0.498

Group # 4 -- W: 0.295, X: 0.899

Group # 8 -- M: 0.246, S: 0.896

Log likelihood = -16.817 Significance = 0.998

Run # 25, 14 cells:

Convergence at Iteration 9

Input 0.004

Group # 2 -- 1: 0.413, 2: 0.940

Group # 4 -- W: 0.305, X: 0.888

Group # 5 -- P: 0.480, A: 0.720

Group # 8 -- M: 0.250, S: 0.893

Log likelihood = -16.503 Significance = 0.444

Run # 26, 16 cells:

Convergence at Iteration 14

Input 0.003

Group # 2 -- 1: 0.412, 2: 0.940

Group # 4 -- W: 0.279, X: 0.915

Group # 6 -- H: 0.597, M: 0.383

Group # 8 -- M: 0.224, S: 0.916

Log likelihood = -16.456 Significance = 0.414

Run # 27, 16 cells:

Convergence at Iteration 13

Input 0.003

Group # 2 -- 1: 0.412, 2: 0.941

Group # 4 -- W: 0.279, X: 0.915

Group # 7 -- D: 0.580, T: 0.373

Group # 8 -- M: 0.222, S: 0.918

Log likelihood = -16.486 Significance = 0.433

No remaining groups significant

Groups selected while stepping up:

2 (características sintático-semânticas do sujeito da oração principal)

8 (nível de escolarização)

4 (explicitação do sujeito da adverbial final)

Best stepping up run: #19

Stepping down...

----- Level # 8 -----

Run # 28, 56 cells:

No Convergence at Iteration 20

Input 0.002

Group # 1 -- =: 0.325, &: 0.823, #: 0.816

Group # 2 -- 1: 0.419, 2: 0.926

Group # 3 -- 3: 0.498, 4: 0.538

Group # 4 -- W: 0.335, X: 0.848

Group # 5 -- P: 0.482, A: 0.705

Group # 6 -- H: 0.715, M: 0.246

Group # 7 -- D: 0.477, T: 0.537

Group # 8 -- M: 0.197, S: 0.937

Log likelihood = -14.699

----- Level # 7 -----

Run # 29, 34 cells:

No Convergence at Iteration 20

Input 0.002

Group # 2 -- 1: 0.407, 2: 0.949

Group # 3 -- 3: 0.492, 4: 0.625

Group # 4 -- W: 0.286, X: 0.908

Group # 5 -- P: 0.476, A: 0.761

Group # 6 -- H: 0.656, M: 0.313

Group # 7 -- D: 0.466, T: 0.554

Group # 8 -- M: 0.225, S: 0.915

Log likelihood = -15.975 Significance = 0.282

Run # 30, 43 cells:

No Convergence at Iteration 20

Input 0.003

Group # 1 -- =: 0.281, &: 0.740, #: 0.920

Group # 3 -- 3: 0.497, 4: 0.547

Group # 4 -- W: 0.357, X: 0.813

Group # 5 -- P: 0.486, A: 0.665

Group # 6 -- H: 0.733, M: 0.226

Group # 7 -- D: 0.431, T: 0.610

Group # 8 -- M: 0.217, S: 0.922

Log likelihood = -18.027 Significance = 0.010

Run # 31, 48 cells:

No Convergence at Iteration 20

Input 0.002

Group # 1 -- =: 0.324, &: 0.823, #: 0.818

Group # 2 -- 1: 0.419, 2: 0.926

Group # 4 -- W: 0.336, X: 0.847

Group # 5 -- P: 0.482, A: 0.705

Group # 6 -- H: 0.714, M: 0.247

Group # 7 -- D: 0.477, T: 0.538

Group # 8 -- M: 0.198, S: 0.937

Log likelihood = -14.704 Significance = 0.928

Run # 32, 41 cells:

No Convergence at Iteration 20

Input 0.002
Group # 1 -- =: 0.244, &: 0.863, #: 0.932
Group # 2 -- 1: 0.428, 2: 0.906
Group # 3 -- 3: 0.503, 4: 0.446
Group # 5 -- P: 0.476, A: 0.754
Group # 6 -- H: 0.714, M: 0.247
Group # 7 -- D: 0.476, T: 0.538
Group # 8 -- M: 0.184, S: 0.946
Log likelihood = -16.595 Significance = 0.052

Run # 33, 46 cells:
No Convergence at Iteration 20
Input 0.002
Group # 1 -- =: 0.313, &: 0.859, #: 0.820
Group # 2 -- 1: 0.420, 2: 0.924
Group # 3 -- 3: 0.498, 4: 0.530
Group # 4 -- W: 0.329, X: 0.857
Group # 6 -- H: 0.709, M: 0.252
Group # 7 -- D: 0.484, T: 0.526
Group # 8 -- M: 0.191, S: 0.941
Log likelihood = -14.912 Significance = 0.519

Run # 34, 51 cells:
Convergence at Iteration 18
Input 0.002
Group # 1 -- =: 0.326, &: 0.822, #: 0.814
Group # 2 -- 1: 0.418, 2: 0.929
Group # 3 -- 3: 0.497, 4: 0.542
Group # 4 -- W: 0.333, X: 0.851
Group # 5 -- P: 0.482, A: 0.698
Group # 7 -- D: 0.666, T: 0.248
Group # 8 -- M: 0.192, S: 0.941
Log likelihood = -14.780 Significance = 0.691

Run # 35, 51 cells:
Convergence at Iteration 17
Input 0.002
Group # 1 -- =: 0.324, &: 0.823, #: 0.817
Group # 2 -- 1: 0.419, 2: 0.926
Group # 3 -- 3: 0.498, 4: 0.539
Group # 4 -- W: 0.335, X: 0.848
Group # 5 -- P: 0.482, A: 0.704
Group # 6 -- H: 0.694, M: 0.269
Group # 8 -- M: 0.196, S: 0.938
Log likelihood = -14.702 Significance = 0.938

Run # 36, 41 cells:
No Convergence at Iteration 20
Input 0.006
Group # 1 -- =: 0.372, &: 0.801, #: 0.710
Group # 2 -- 1: 0.428, 2: 0.906
Group # 3 -- 3: 0.497, 4: 0.552
Group # 4 -- W: 0.329, X: 0.857
Group # 5 -- P: 0.474, A: 0.773
Group # 6 -- H: 0.700, M: 0.263
Group # 7 -- D: 0.320, T: 0.771
Log likelihood = -20.505 Significance = 0.001

Cut Group # 7 with factors DT

----- Level # 6 -----

Run # 37, 30 cells:

Convergence at Iteration 15

Input 0.002

Group # 2 -- 1: 0.407, 2: 0.950

Group # 3 -- 3: 0.492, 4: 0.627

Group # 4 -- W: 0.286, X: 0.909

Group # 5 -- P: 0.476, A: 0.761

Group # 6 -- H: 0.621, M: 0.354

Group # 8 -- M: 0.224, S: 0.916

Log likelihood = -15.983 Significance = 0.281

Run # 38, 38 cells:

Convergence at Iteration 14

Input 0.003

Group # 1 -- =: 0.279, &: 0.743, #: 0.923

Group # 3 -- 3: 0.497, 4: 0.550

Group # 4 -- W: 0.358, X: 0.813

Group # 5 -- P: 0.486, A: 0.662

Group # 6 -- H: 0.672, M: 0.294

Group # 8 -- M: 0.212, S: 0.926

Log likelihood = -18.061 Significance = 0.010

Run # 39, 43 cells:

Convergence at Iteration 18

Input 0.002

Group # 1 -- =: 0.324, &: 0.823, #: 0.820

Group # 2 -- 1: 0.419, 2: 0.926

Group # 4 -- W: 0.336, X: 0.847

Group # 5 -- P: 0.482, A: 0.705

Group # 6 -- H: 0.693, M: 0.271

Group # 8 -- M: 0.197, S: 0.938

Log likelihood = -14.707 Significance = 0.927

Run # 40, 37 cells:

No Convergence at Iteration 20

Input 0.002

Group # 1 -- =: 0.243, &: 0.863, #: 0.933

Group # 2 -- 1: 0.428, 2: 0.906

Group # 3 -- 3: 0.503, 4: 0.447

Group # 5 -- P: 0.476, A: 0.753

Group # 6 -- H: 0.693, M: 0.270

Group # 8 -- M: 0.183, S: 0.947

Log likelihood = -16.599 Significance = 0.052

Run # 41, 42 cells:

Convergence at Iteration 18

Input 0.002

Group # 1 -- =: 0.312, &: 0.859, #: 0.821

Group # 2 -- 1: 0.420, 2: 0.925

Group # 3 -- 3: 0.498, 4: 0.531

Group # 4 -- W: 0.329, X: 0.857

Group # 6 -- H: 0.695, M: 0.268

Group # 8 -- M: 0.191, S: 0.942

Log likelihood = -14.914 Significance = 0.521

Run # 42, 35 cells:
 Convergence at Iteration 11
 Input 0.003
 Group # 1 -- =: 0.382, &: 0.793, #: 0.688
 Group # 2 -- 1: 0.420, 2: 0.924
 Group # 3 -- 3: 0.502, 4: 0.475
 Group # 4 -- W: 0.338, X: 0.843
 Group # 5 -- P: 0.483, A: 0.695
 Group # 8 -- M: 0.252, S: 0.891
 Log likelihood = -15.746 Significance = 0.158

Run # 43, 35 cells:
 Convergence at Iteration 11
 Input 0.006
 Group # 1 -- =: 0.375, &: 0.799, #: 0.703
 Group # 2 -- 1: 0.426, 2: 0.912
 Group # 3 -- 3: 0.495, 4: 0.581
 Group # 4 -- W: 0.321, X: 0.867
 Group # 5 -- P: 0.476, A: 0.755
 Group # 6 -- H: 0.511, M: 0.486
 Log likelihood = -20.968 Significance = 0.000

Cut Group # 3 with factors 34

----- Level # 5 -----

Run # 44, 24 cells:
 Convergence at Iteration 15
 Input 0.002
 Group # 2 -- 1: 0.406, 2: 0.950
 Group # 4 -- W: 0.286, X: 0.908
 Group # 5 -- P: 0.476, A: 0.756
 Group # 6 -- H: 0.614, M: 0.363
 Group # 8 -- M: 0.224, S: 0.916
 Log likelihood = -16.025 Significance = 0.272

Run # 45, 31 cells:
 Convergence at Iteration 13
 Input 0.003
 Group # 1 -- =: 0.278, &: 0.742, #: 0.924
 Group # 4 -- W: 0.357, X: 0.814
 Group # 5 -- P: 0.486, A: 0.657
 Group # 6 -- H: 0.670, M: 0.297
 Group # 8 -- M: 0.212, S: 0.926
 Log likelihood = -18.071 Significance = 0.010

Run # 46, 29 cells:
 No Convergence at Iteration 20
 Input 0.002
 Group # 1 -- =: 0.244, &: 0.864, #: 0.932
 Group # 2 -- 1: 0.428, 2: 0.904
 Group # 5 -- P: 0.477, A: 0.748
 Group # 6 -- H: 0.694, M: 0.270
 Group # 8 -- M: 0.180, S: 0.948
 Log likelihood = -16.610 Significance = 0.052

Run # 47, 34 cells:
 Convergence at Iteration 18
 Input 0.002

Group # 1 -- =: 0.312, &: 0.858, #: 0.823
Group # 2 -- 1: 0.420, 2: 0.924
Group # 4 -- W: 0.329, X: 0.856
Group # 6 -- H: 0.694, M: 0.270
Group # 8 -- M: 0.191, S: 0.942
Log likelihood = -14.916 Significance = 0.524

Run # 48, 29 cells:

Convergence at Iteration 11

Input 0.003

Group # 1 -- =: 0.383, &: 0.794, #: 0.685
Group # 2 -- 1: 0.420, 2: 0.924
Group # 4 -- W: 0.338, X: 0.843
Group # 5 -- P: 0.483, A: 0.696
Group # 8 -- M: 0.252, S: 0.890
Log likelihood = -15.747 Significance = 0.159

Run # 49, 28 cells:

Convergence at Iteration 10

Input 0.007

Group # 1 -- =: 0.372, &: 0.799, #: 0.712
Group # 2 -- 1: 0.425, 2: 0.913
Group # 4 -- W: 0.327, X: 0.860
Group # 5 -- P: 0.476, A: 0.761
Group # 6 -- H: 0.511, M: 0.486
Log likelihood = -20.996 Significance = 0.000

Cut Group # 5 with factors PA

----- Level # 4 -----

Run # 50, 16 cells:

Convergence at Iteration 14

Input 0.003

Group # 2 -- 1: 0.412, 2: 0.940
Group # 4 -- W: 0.279, X: 0.915
Group # 6 -- H: 0.597, M: 0.383
Group # 8 -- M: 0.224, S: 0.916
Log likelihood = -16.456 Significance = 0.217

Run # 51, 23 cells:

Convergence at Iteration 15

Input 0.003

Group # 1 -- =: 0.272, &: 0.776, #: 0.923
Group # 4 -- W: 0.349, X: 0.827
Group # 6 -- H: 0.668, M: 0.299
Group # 8 -- M: 0.206, S: 0.930
Log likelihood = -18.227 Significance = 0.010

Run # 52, 20 cells:

No Convergence at Iteration 20

Input 0.002

Group # 1 -- =: 0.228, &: 0.895, #: 0.939
Group # 2 -- 1: 0.429, 2: 0.903
Group # 6 -- H: 0.702, M: 0.260
Group # 8 -- M: 0.178, S: 0.950
Log likelihood = -17.047 Significance = 0.042

Run # 53, 20 cells:

Convergence at Iteration 11
Input 0.003
Group # 1 -- =: 0.380, &: 0.814, #: 0.678
Group # 2 -- 1: 0.423, 2: 0.919
Group # 4 -- W: 0.328, X: 0.858
Group # 8 -- M: 0.246, S: 0.896
Log likelihood = -15.968 Significance = 0.157

Run # 54, 20 cells:
Convergence at Iteration 10
Input 0.008
Group # 1 -- =: 0.384, &: 0.780, #: 0.690
Group # 2 -- 1: 0.429, 2: 0.904
Group # 4 -- W: 0.324, X: 0.863
Group # 6 -- H: 0.500, M: 0.500
Log likelihood = -21.634 Significance = 0.000

Cut Group # 1 with factors =&#

----- Level # 3 -----

Run # 55, 8 cells:
Convergence at Iteration 7
Input 0.007
Group # 4 -- W: 0.273, X: 0.921
Group # 6 -- H: 0.520, M: 0.476
Group # 8 -- M: 0.254, S: 0.888
Log likelihood = -21.977 Significance = 0.001

Run # 56, 8 cells:
Convergence at Iteration 7
Input 0.009
Group # 2 -- 1: 0.407, 2: 0.950
Group # 6 -- H: 0.500, M: 0.500
Group # 8 -- M: 0.255, S: 0.887
Log likelihood = -21.780 Significance = 0.001

Run # 57, 8 cells:
Convergence at Iteration 8
Input 0.004
Group # 2 -- 1: 0.417, 2: 0.931
Group # 4 -- W: 0.295, X: 0.899
Group # 8 -- M: 0.247, S: 0.896
Log likelihood = -16.817 Significance = 0.414

Run # 58, 8 cells:
Convergence at Iteration 6
Input 0.011
Group # 2 -- 1: 0.422, 2: 0.920
Group # 4 -- W: 0.307, X: 0.886
Group # 6 -- H: 0.473, M: 0.533
Log likelihood = -22.492 Significance = 0.001

Cut Group # 6 with factors HM

----- Level # 2 -----

Run # 59, 4 cells:
Convergence at Iteration 7

Input 0.007
Group # 4 -- W: 0.276, X: 0.919
Group # 8 -- M: 0.256, S: 0.886
Log likelihood = -22.001 Significance = 0.002

Run # 60, 4 cells:
Convergence at Iteration 7
Input 0.009
Group # 2 -- 1: 0.407, 2: 0.950
Group # 8 -- M: 0.255, S: 0.887
Log likelihood = -21.780 Significance = 0.003

Run # 61, 4 cells:
Convergence at Iteration 6
Input 0.011
Group # 2 -- 1: 0.422, 2: 0.920
Group # 4 -- W: 0.305, X: 0.888
Log likelihood = -22.534 Significance = 0.001

All remaining groups significant

Groups eliminated while stepping down: 7 3 5 1 6
Best stepping up run: #19
Best stepping down run: #57

Anexo 4: Rodada dos dados da escrita

• GROUPS & FACTORS • 7/7/2011 22:21:08


```

-----
Group      Default    Factors
  1         D        DR
  2         &        &=#
  3         2        21
  4         3        34
  5         X        XW
  6         P        PA
  7         M        MH
  8         D        DTQ
  
```

• CELL CREATION • 7/7/2011 22:22:22


```

      Name of token file: Dados da escrita.tkn
Name of condition file: Condições escrita.cnd
(
; Identity recode: All groups included as is.
(1)
(2)
(3)
(4)
(5)
(6)
(7)
(8)
)
  
```

```

      Number of cells: 75
Application value(s): DR
Total no. of factors: 16
  
```

Group		D	R	Total	%	

1 (2)		D	R			
& N		8	19	27	9.6	sujeito parcialmente co
rreferente						
%		29.6	70.4			
= N		4	187	191	67.7	sujeito correferente
%		2.1	97.9			
# N		33	31	64	22.7	sujeito não correferent
e						
%		51.6	48.4			
Total N		45	237	282		
%		16.0	84.0			

2 (3)		D	R			
2 N		27	76	103	36.5	sujeito não controlador
%		26.2	73.8			

1	N	18	161	179	63.5	sujeito controlador
	%	10.1	89.9			
Total	N	45	237	282		
	%	16.0	84.0			
<hr/>						
3 (4)		D	R			
3	N	37	223	260	92.2	voz ativa
	%	14.2	85.8			
4	N	8	14	22	7.8	voz passiva
	%	36.4	63.6			
Total	N	45	237	282		
	%	16.0	84.0			
<hr/>						
4 (5)		D	R			
X	N	36	6	42	14.9	sujeito explícito
	%	85.7	14.3			
W	N	9	231	240	85.1	sujeito não explícito
	%	3.8	96.2			
Total	N	45	237	282		
	%	16.0	84.0			
<hr/>						
5 (6)		D	R			
P	N	36	170	206	73.0	final posposta
	%	17.5	82.5			
A	N	9	67	76	27.0	final anteposta
	%	11.8	88.2			
Total	N	45	237	282		
	%	16.0	84.0			
<hr/>						
6 (7)		D	R			
M	N	15	68	83	29.4	mulheres
	%	18.1	81.9			
H	N	30	169	199	70.6	homens
	%	15.1	84.9			
Total	N	45	237	282		
	%	16.0	84.0			
<hr/>						
7 (8)		D	R			
D	N	12	68	80	28.4	15-25 anos
	%	15.0	85.0			
T	N	22	117	139	49.3	26-49 anos
	%	15.8	84.2			
Q	N	11	52	63	22.3	50 anos ou mais
	%	17.5	82.5			
Total	N	45	237	282		
	%	16.0	84.0			
<hr/>						
TOTAL	N	45	237	282		

% 16.0 84.0

Name of new cell file: Condições escrita.cel

• BINOMIAL VARBRUL • 7/7/2011 22:23:11
.....

Name of cell file: Condições escrita.cel

Averaging by weighting factors.

Threshold, step-up/down: 0.050001

Stepping up...

----- Level # 0 -----

Run # 1, 1 cells:

Convergence at Iteration 2

Input 0.160

Log likelihood = -123.788

----- Level # 1 -----

Run # 2, 3 cells:

Convergence at Iteration 6

Input 0.065

Group # 1 -- &: 0.859, =: 0.237, #: 0.939

Log likelihood = -80.160 Significance = 0.000

Run # 3, 2 cells:

Convergence at Iteration 5

Input 0.146

Group # 2 -- 2: 0.675, 1: 0.396

Log likelihood = -117.663 Significance = 0.000

Run # 4, 2 cells:

Convergence at Iteration 5

Input 0.155

Group # 3 -- 3: 0.476, 4: 0.757

Log likelihood = -120.795 Significance = 0.015

Run # 5, 2 cells:

Convergence at Iteration 6

Input 0.076

Group # 4 -- X: 0.986, W: 0.321

Log likelihood = -55.605 Significance = 0.000

Run # 6, 2 cells:

Convergence at Iteration 4

Input 0.158

Group # 5 -- P: 0.531, A: 0.418

Log likelihood = -123.096 Significance = 0.246

Run # 7, 2 cells:

Convergence at Iteration 4

Input 0.159

Group # 6 -- M: 0.538, H: 0.484

Log likelihood = -123.595 Significance = 0.546

Run # 8, 3 cells:
Convergence at Iteration 3
Input 0.159
Group # 7 -- D: 0.482, T: 0.498, Q: 0.527
Log likelihood = -123.707 Significance = 0.923

Add Group # 4 with factors XW

----- Level # 2 -----

Run # 9, 6 cells:
Convergence at Iteration 10
Input 0.058
Group # 1 -- &: 0.810, =: 0.355, #: 0.764
Group # 4 -- X: 0.969, W: 0.354
Log likelihood = -51.560 Significance = 0.018

Run # 10, 4 cells:
Convergence at Iteration 6
Input 0.075
Group # 2 -- 2: 0.588, 1: 0.449
Group # 4 -- X: 0.985, W: 0.324
Log likelihood = -55.088 Significance = 0.313

Run # 11, 4 cells:
Convergence at Iteration 6
Input 0.076
Group # 3 -- 3: 0.491, 4: 0.610
Group # 4 -- X: 0.986, W: 0.322
Log likelihood = -55.455 Significance = 0.602

Run # 12, 4 cells:
Convergence at Iteration 8
Input 0.071
Group # 4 -- X: 0.989, W: 0.314
Group # 5 -- P: 0.580, A: 0.295
Log likelihood = -53.970 Significance = 0.075

Run # 13, 4 cells:
Convergence at Iteration 6
Input 0.075
Group # 4 -- X: 0.987, W: 0.320
Group # 6 -- M: 0.573, H: 0.469
Log likelihood = -55.346 Significance = 0.480

Run # 14, 6 cells:
Convergence at Iteration 6
Input 0.076
Group # 4 -- X: 0.987, W: 0.320
Group # 7 -- D: 0.551, T: 0.484, Q: 0.470
Log likelihood = -55.484 Significance = 0.887

Add Group # 1 with factors &=#

----- Level # 3 -----

Run # 15, 11 cells:
Convergence at Iteration 11
Input 0.058
Group # 1 -- &: 0.810, =: 0.359, #: 0.752

Group # 2 -- 2: 0.522, 1: 0.487
Group # 4 -- X: 0.969, W: 0.353
Log likelihood = -51.532 Significance = 0.819

Run # 16, 11 cells:
Convergence at Iteration 10
Input 0.058
Group # 1 -- &: 0.808, =: 0.356, #: 0.762
Group # 3 -- 3: 0.493, 4: 0.579
Group # 4 -- X: 0.968, W: 0.355
Log likelihood = -51.483 Significance = 0.696

Run # 17, 10 cells:
Convergence at Iteration 10
Input 0.056
Group # 1 -- &: 0.773, =: 0.352, #: 0.786
Group # 4 -- X: 0.971, W: 0.351
Group # 5 -- P: 0.578, A: 0.299
Log likelihood = -50.077 Significance = 0.089

Run # 18, 11 cells:
Convergence at Iteration 10
Input 0.056
Group # 1 -- &: 0.807, =: 0.342, #: 0.794
Group # 4 -- X: 0.966, W: 0.357
Group # 6 -- M: 0.622, H: 0.448
Log likelihood = -50.935 Significance = 0.269

Run # 19, 16 cells:
Convergence at Iteration 10
Input 0.057
Group # 1 -- &: 0.808, =: 0.348, #: 0.780
Group # 4 -- X: 0.968, W: 0.355
Group # 7 -- D: 0.586, T: 0.474, Q: 0.448
Log likelihood = -51.249 Significance = 0.735

No remaining groups significant

Groups selected while stepping up:

4 (explicitação do sujeito da adverbial final) 1 (correferencialidade do sujeito da adverbial final)

Best stepping up run: #9

Stepping down...

----- Level # 7 -----

Run # 20, 75 cells:
Convergence at Iteration 12
Input 0.052
Group # 1 -- &: 0.771, =: 0.352, #: 0.788
Group # 2 -- 2: 0.549, 1: 0.472
Group # 3 -- 3: 0.487, 4: 0.647
Group # 4 -- X: 0.970, W: 0.353
Group # 5 -- P: 0.580, A: 0.293
Group # 6 -- M: 0.587, H: 0.463

Group # 7 -- D: 0.566, T: 0.484, Q: 0.452
Log likelihood = -49.162

----- Level # 6 -----

Run # 21, 48 cells:

Convergence at Iteration 8

Input 0.065

Group # 2 -- 2: 0.607, 1: 0.438

Group # 3 -- 3: 0.485, 4: 0.668

Group # 4 -- X: 0.987, W: 0.318

Group # 5 -- P: 0.589, A: 0.275

Group # 6 -- M: 0.548, H: 0.480

Group # 7 -- D: 0.553, T: 0.485, Q: 0.466

Log likelihood = -52.695 Significance = 0.032

Run # 22, 51 cells:

Convergence at Iteration 10

Input 0.053

Group # 1 -- &: 0.770, =: 0.343, #: 0.808

Group # 3 -- 3: 0.489, 4: 0.631

Group # 4 -- X: 0.969, W: 0.354

Group # 5 -- P: 0.577, A: 0.301

Group # 6 -- M: 0.589, H: 0.462

Group # 7 -- D: 0.564, T: 0.484, Q: 0.455

Log likelihood = -49.286 Significance = 0.635

Run # 23, 61 cells:

Convergence at Iteration 12

Input 0.053

Group # 1 -- &: 0.772, =: 0.348, #: 0.795

Group # 2 -- 2: 0.535, 1: 0.480

Group # 4 -- X: 0.970, W: 0.352

Group # 5 -- P: 0.578, A: 0.300

Group # 6 -- M: 0.581, H: 0.466

Group # 7 -- D: 0.552, T: 0.488, Q: 0.459

Log likelihood = -49.433 Significance = 0.472

Run # 24, 63 cells:

Convergence at Iteration 8

Input 0.057

Group # 1 -- &: 0.820, =: 0.221, #: 0.958

Group # 2 -- 2: 0.479, 1: 0.512

Group # 3 -- 3: 0.480, 4: 0.720

Group # 5 -- P: 0.569, A: 0.319

Group # 6 -- M: 0.659, H: 0.432

Group # 7 -- D: 0.524, T: 0.498, Q: 0.475

Log likelihood = -74.155 Significance = 0.000

Run # 25, 57 cells:

Convergence at Iteration 11

Input 0.055

Group # 1 -- &: 0.804, =: 0.347, #: 0.784

Group # 2 -- 2: 0.527, 1: 0.484

Group # 3 -- 3: 0.488, 4: 0.637

Group # 4 -- X: 0.966, W: 0.358

Group # 6 -- M: 0.611, H: 0.453

Group # 7 -- D: 0.547, T: 0.484, Q: 0.477

Log likelihood = -50.654 Significance = 0.088

Run # 26, 55 cells:

Convergence at Iteration 11

Input 0.053

Group # 1 -- &: 0.771, =: 0.357, #: 0.775

Group # 2 -- 2: 0.552, 1: 0.470

Group # 3 -- 3: 0.488, 4: 0.638

Group # 4 -- X: 0.971, W: 0.351

Group # 5 -- P: 0.585, A: 0.283

Group # 7 -- D: 0.606, T: 0.472, Q: 0.426

Log likelihood = -49.396 Significance = 0.495

Run # 27, 43 cells:

Convergence at Iteration 12

Input 0.053

Group # 1 -- &: 0.772, =: 0.353, #: 0.785

Group # 2 -- 2: 0.548, 1: 0.472

Group # 3 -- 3: 0.489, 4: 0.631

Group # 4 -- X: 0.969, W: 0.353

Group # 5 -- P: 0.577, A: 0.302

Group # 6 -- M: 0.619, H: 0.450

Log likelihood = -49.313 Significance = 0.861

Cut Group # 7 with factors DTQ

----- Level # 5 -----

Run # 28, 25 cells:

Convergence at Iteration 8

Input 0.066

Group # 2 -- 2: 0.607, 1: 0.438

Group # 3 -- 3: 0.486, 4: 0.658

Group # 4 -- X: 0.987, W: 0.319

Group # 5 -- P: 0.585, A: 0.282

Group # 6 -- M: 0.572, H: 0.470

Log likelihood = -52.796 Significance = 0.034

Run # 29, 27 cells:

Convergence at Iteration 10

Input 0.054

Group # 1 -- &: 0.770, =: 0.344, #: 0.805

Group # 3 -- 3: 0.490, 4: 0.615

Group # 4 -- X: 0.968, W: 0.355

Group # 5 -- P: 0.574, A: 0.309

Group # 6 -- M: 0.620, H: 0.449

Log likelihood = -49.428 Significance = 0.647

Run # 30, 30 cells:

Convergence at Iteration 12

Input 0.054

Group # 1 -- &: 0.773, =: 0.349, #: 0.793

Group # 2 -- 2: 0.535, 1: 0.480

Group # 4 -- X: 0.970, W: 0.352

Group # 5 -- P: 0.575, A: 0.305

Group # 6 -- M: 0.607, H: 0.455

Log likelihood = -49.531 Significance = 0.513

Run # 31, 31 cells:

Convergence at Iteration 9
Input 0.057
Group # 1 -- &: 0.821, =: 0.221, #: 0.957
Group # 2 -- 2: 0.478, 1: 0.512
Group # 3 -- 3: 0.480, 4: 0.716
Group # 5 -- P: 0.569, A: 0.320
Group # 6 -- M: 0.670, H: 0.427
Log likelihood = -74.201 Significance = 0.000

Run # 32, 30 cells:
Convergence at Iteration 12
Input 0.055
Group # 1 -- &: 0.805, =: 0.348, #: 0.781
Group # 2 -- 2: 0.528, 1: 0.484
Group # 3 -- 3: 0.489, 4: 0.627
Group # 4 -- X: 0.966, W: 0.358
Group # 6 -- M: 0.632, H: 0.444
Log likelihood = -50.721 Significance = 0.095

Run # 33, 29 cells:
Convergence at Iteration 11
Input 0.055
Group # 1 -- &: 0.772, =: 0.364, #: 0.760
Group # 2 -- 2: 0.549, 1: 0.472
Group # 3 -- 3: 0.492, 4: 0.596
Group # 4 -- X: 0.971, W: 0.351
Group # 5 -- P: 0.581, A: 0.292
Log likelihood = -49.876 Significance = 0.291

Cut Group # 2 with factors 21

----- Level # 4 -----

Run # 34, 14 cells:
Convergence at Iteration 7
Input 0.070
Group # 3 -- 3: 0.487, 4: 0.646
Group # 4 -- X: 0.988, W: 0.316
Group # 5 -- P: 0.579, A: 0.296
Group # 6 -- M: 0.568, H: 0.471
Log likelihood = -53.549 Significance = 0.017

Run # 35, 18 cells:
Convergence at Iteration 10
Input 0.054
Group # 1 -- &: 0.772, =: 0.342, #: 0.807
Group # 4 -- X: 0.969, W: 0.354
Group # 5 -- P: 0.574, A: 0.309
Group # 6 -- M: 0.609, H: 0.454
Log likelihood = -49.594 Significance = 0.580

Run # 36, 18 cells:
Convergence at Iteration 8
Input 0.056
Group # 1 -- &: 0.822, =: 0.224, #: 0.955
Group # 3 -- 3: 0.480, 4: 0.722
Group # 5 -- P: 0.570, A: 0.317
Group # 6 -- M: 0.669, H: 0.427
Log likelihood = -74.242 Significance = 0.000

Run # 37, 18 cells:
Convergence at Iteration 10
Input 0.055
Group # 1 -- &: 0.804, =: 0.343, #: 0.794
Group # 3 -- 3: 0.490, 4: 0.620
Group # 4 -- X: 0.965, W: 0.359
Group # 6 -- M: 0.632, H: 0.444
Log likelihood = -50.762 Significance = 0.104

Run # 38, 17 cells:
Convergence at Iteration 10
Input 0.056
Group # 1 -- &: 0.771, =: 0.354, #: 0.782
Group # 3 -- 3: 0.493, 4: 0.578
Group # 4 -- X: 0.970, W: 0.352
Group # 5 -- P: 0.578, A: 0.298
Log likelihood = -49.997 Significance = 0.289

Cut Group # 3 with factors 34

----- Level # 3 -----

Run # 39, 8 cells:
Convergence at Iteration 7
Input 0.071
Group # 4 -- X: 0.989, W: 0.314
Group # 5 -- P: 0.577, A: 0.300
Group # 6 -- M: 0.556, H: 0.477
Log likelihood = -53.825 Significance = 0.015

Run # 40, 11 cells:
Convergence at Iteration 8
Input 0.058
Group # 1 -- &: 0.827, =: 0.222, #: 0.956
Group # 5 -- P: 0.572, A: 0.314
Group # 6 -- M: 0.647, H: 0.437
Log likelihood = -75.541 Significance = 0.000

Run # 41, 11 cells:
Convergence at Iteration 10
Input 0.056
Group # 1 -- &: 0.807, =: 0.342, #: 0.794
Group # 4 -- X: 0.966, W: 0.357
Group # 6 -- M: 0.622, H: 0.448
Log likelihood = -50.935 Significance = 0.102

Run # 42, 10 cells:
Convergence at Iteration 10
Input 0.056
Group # 1 -- &: 0.773, =: 0.352, #: 0.786
Group # 4 -- X: 0.971, W: 0.351
Group # 5 -- P: 0.578, A: 0.299
Log likelihood = -50.077 Significance = 0.335

Cut Group # 6 with factors MH

----- Level # 2 -----

Run # 43, 4 cells:

Convergence at Iteration 8
Input 0.071
Group # 4 -- X: 0.989, W: 0.314
Group # 5 -- P: 0.580, A: 0.295
Log likelihood = -53.970 Significance = 0.021

Run # 44, 6 cells:
Convergence at Iteration 7
Input 0.060
Group # 1 -- &: 0.831, =: 0.229, #: 0.950
Group # 5 -- P: 0.579, A: 0.296
Log likelihood = -77.143 Significance = 0.000

Run # 45, 6 cells:
Convergence at Iteration 10
Input 0.058
Group # 1 -- &: 0.810, =: 0.355, #: 0.764
Group # 4 -- X: 0.969, W: 0.354
Log likelihood = -51.560 Significance = 0.089

Cut Group # 5 with factors PA

----- Level # 1 -----

Run # 46, 2 cells:
Convergence at Iteration 6
Input 0.076
Group # 4 -- X: 0.986, W: 0.321
Log likelihood = -55.605 Significance = 0.018

Run # 47, 3 cells:
Convergence at Iteration 6
Input 0.065
Group # 1 -- &: 0.859, =: 0.237, #: 0.939
Log likelihood = -80.160 Significance = 0.000

All remaining groups significant

Groups eliminated while stepping down: 7 2 3 6 5
Best stepping up run: #9
Best stepping down run: #45